



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : C07K 7/00, A61K 39/12 C12Q 1/70	A1	(11) International Publication Number: WO 94/06826 (43) International Publication Date: 31 March 1994 (31.03.94)
(21) International Application Number: PCT/US93/08638 (22) International Filing Date: 15 September 1993 (15.09.93) (30) Priority data: 946,054 15 September 1992 (15.09.92) US (71) Applicant: UNITED BIOMEDICAL, INC. [US/US]; 25 Davids Drive, Hauppauge, NY 11788 (US). (72) Inventors: WANG, Chang, Yi ; 47 Snake Hill Road, Cold Spring Harbor, NY 11724 (US). HOSEIN, Barbara ; 196 East 75th Street, New York, NY 10021 (US). (74) Agent: WILSON, M., Lisa; United Biomedical, Inc., 25 Davids Drive, Hauppauge, NY 11788 (US).		(81) Designated States: AU, CA, FI, JP, NO, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report</i>
(54) Title: NOVEL BRANCHED HYBRID AND CLUSTER PEPTIDES EFFECTIVE IN DIAGNOSING AND DETECTING NON-A, NON-B HEPATITIS (57) Abstract The present invention relates to novel branched peptides specific for the diagnosis and prevention of non-A, non-B hepatitis (NANBH), as well as hepatitis C virus (HCV) infection. More particularly, the present invention is directed to branched synthetic substituted and hybrid peptides containing at least one epitope which is effective in detecting NANBH-associated antibodies in patients with NANBH using immunoassay techniques. In addition, this invention provides immunoassays and kits for the detection and diagnosis of NANBH or HCV infection using the subject peptides.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NE	Niger
BE	Belgium	GN	Guinea	NL	Netherlands
BF	Burkina Faso	GR	Greece	NO	Norway
BG	Bulgaria	HU	Hungary	NZ	New Zealand
BJ	Benin	IE	Ireland	PL	Poland
BR	Brazil	IT	Italy	PT	Portugal
BY	Belarus	JP	Japan	RO	Romania
CA	Canada	KP	Democratic People's Republic of Korea	RU	Russian Federation
CF	Central African Republic	KR	Republic of Korea	SD	Sudan
CG	Congo	KZ	Kazakhstan	SE	Sweden
CH	Switzerland	LJ	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovak Republic
CM	Cameroon	LU	Luxembourg	SN	Senegal
CN	China	LV	Latvia	TD	Chad
CS	Czechoslovakia	MC	Monaco	TG	Togo
CZ	Czech Republic	MG	Madagascar	UA	Ukraine
DE	Germany	ML	Mali	US	United States of America
DK	Denmark	MN	Mongolia	UZ	Uzbekistan
ES	Spain			VN	Viet Nam
FI	Finland				

1 NOVEL BRANCHED HYBRID AND CLUSTER PEPTIDES EFFECTIVE IN
DIAGNOSING AND DETECTING NON-A, NON-B HEPATITIS

The present invention relates to novel
branching peptides specific for the diagnosis and
5 prevention of non-A, non-B hepatitis (NANBH), including
hepatitis C virus (HCV) infection. More particularly,
the present invention is directed to branched synthetic
peptides containing at least one epitope which is
effective in detecting NANBH-associated antibodies in
10 patients with NANBH using immunoassay techniques.
Further, the present invention is directed to synthetic
peptides which are hybrids of the peptides described
herein.

Non-A, non-B hepatitis (NANBH) remains the
15 most common form of post-transfusion hepatitis, imposing
a strong need for sensitive and specific diagnostic
screening methods to identify potential blood donors and
other persons who may be carriers of the disease. Thus,
accurate screening methods are needed to permit removal
20 of contaminated blood and blood products from the blood
supply with a high degree confidence.

The etiological agent of NANBH, HCV, has been
cloned and identified by several groups [Houghton et
al., EP 0318216, published 5/1989; Okamoto et al.
25 (1990) Jpn. J. Exp. Med. 60:167; Houghton et al., EP
0388232, published 9/1990; and Kato et al. (1990) Proc.
Natl. Acad. Sci. USA 87:9524; Arima et al. (1989a)
Gastroenterologia Japonica 24:540; Reyes et al. (1990)
Science 247:1335; Arima et al. (1989b) Gastroenterologia
30 Japonica 24:545; Maeno et al. (1990) Nucleic Acids Res.
18:2685]. The HCV genome is about 10 kilobases (kb) in
length and encodes a single polyprotein which is

1 processed into structural and non-structural proteins.
From the N terminus, the polyprotein includes the capsid
and envelope proteins of the structural region and the
NS-1 to NS-5 proteins of the non-structural region.

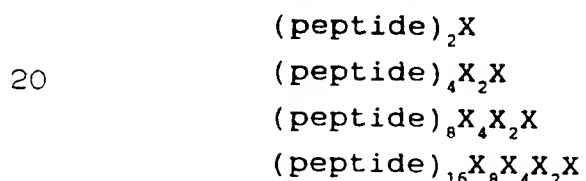
5 While some of the antigenic regions of HCV
have been identified, peptides and recombinant proteins
from these regions exhibit a variable degree of
sensitivity and selectivity in detection and diagnosis
of NANBH carriers. Antigenic regions have been reported
10 in the core, or capsid, protein [Hosein et al. (1991)
Proc. Natl. Acad. Sci. USA 88:3647; UBI HCV EIA Product
Insert (1990); Okamoto et al. (1990) Jap. J. Exp. Med.
60:223; U.S. Patent No. 5,106,726; Takahashi et al.
(1992) J. Gen. Virol. 73:667; Kotwal et al. (1992) Proc.
15 Natl. Acad. Sci. USA 89:4486]; in the envelope, NS-1,
NS-2 and NS-3 proteins [Wang et al., EP 0468527,
published Jan. 29, 1992]; NS-4 protein [Houghton (1989);
Kuo et al. (1989) Science 244:362; U.S. Patent No.
5,106,726] and NS-5 protein [Maeno et al. (1990) Nucleic
20 Acids Res. 18:2685; Wang (1992)].

In addition to HCV-derived antigens, there
exist other NANBH-associated antigens that appear to be
encoded by a host cellular sequence. One such antigen,
known as the GOR epitope, is reactive with sera from
25 individuals who are PCR positive for HCV [Mishiro et al.
(1990) Lancet 336:1400].

Serological validation has been used to map
epitopes within certain HCV antigenic regions as
described in Wang (1992) and U.S. Patent No. 5,106,726,
30 each of which is incorporated herein by reference.
These mapping studies employed synthetic peptides to
screen well-characterized NANBH serum panels and

1 permitted identification of strong HCV antigens.
 Further refinement of the epitope analysis using
 serological validation techniques has led to the
 discovery that small clusters of amino acid residues
 5 contained within longer branched peptides or fusions of
 peptides containing one or more epitopes from separate
 regions of the HCV genome provide a superior and more
 sensitive assay for diagnosis and detection of NANBH
 carriers as well as for HCV infection. Hence, the
 10 present invention permits earlier detection of NANBH
 seroconversion and shows improved specificity, for
 example, fewer false positive serum samples are
 detected.

The present invention relates to branched
 15 synthetic peptides for the diagnosis and detection of
 NANBH and HCV infection. In particular the subject
 peptides are provided as a peptide composition having at
 least one branched peptide represented by the formula



where X is an amino acid or an amino acid analog having
 two amino groups and one carboxyl group with each group
 25 being capable of forming a peptide bond linkage, and
 where the peptide moiety comprises at least one epitope
 which is specifically immunoreactive with antibodies
 against HCV. The peptide moiety further comprises at
 least one cluster of from about 3 to about 20 contiguous
 30 amino acids from the sequences:

- 1 Gly-Cys-Ser-Gly-Gly-Ala-Tyr-Asp-Ile-Ile-Ile-Cys-Asp-Glu-Leu-
His-Ser-Thr-Asp-Ala-Thr-Ser-Ile-Leu-Gly-Ile-Gly-Thr-Val-Leu-
Asp-Gln-Ala-Glu-Thr-Ala-Gly, (Pep3; SEQ ID NO:1),
- 5 Phe-Thr-Phe-Ser-Pro-Arg-Arg-His-Trp-Thr-Thr-Gln-Gly-Cys-Asn-
Cys-Ser-Ile-Tyr-Pro-Gly-His-Ile-Thr-Gly-His-Arg-Met-Ala-Trp-
Asp-Met-Met-Met-Asn-Trp-Ser-Pro-Thr-Ala, (Pep8; SEQ ID
NO:2),
- 10 Glu-Ile-Leu-Arg-Lys-Ser-Arg-Arg-Phe-Ala-Gln-Ala-Leu-Pro-Val-
Trp-Ala-Arg-Pro-Asp-Tyr-Asn-Pro-Pro-Leu-Val-Glu-Thr-Trp-Lys-
Lys-Pro-Asp-Tyr-Glu-Pro-Pro-Val-Val-His-Gly-Cys-Pro-Leu-Pro-
Pro-Pro-Lys-Ser-Pro-Pro-Val-Pro-Pro-Pro-Arg-Lys-Lys-Arg-Thr,
(Pep11; SEQ ID NO:3),
- 15 Glu-Ile-Pro-Phe-Tyr-Gly-Lys-Ala-Ile-Pro-Leu-Glu-Val-Ile-Lys-
Gly-Gly-Arg-His-Leu-Ile-Phe-Cys-His-Ser-Lys-Lys-Lys-Cys-Asp-
Glu-Leu-Ala-Ala-Lys-Leu-Val-Ala-Leu, (Pep18; SEQ ID NO:4),
- 20 Pro-Val-Val-Pro-Gln-Ser-Phe-Gln-Val-Ala-His-Leu-His-Ala-Pro-
Thr-Gly-Ser-Gly-Lys-Ser, (Pep25; SEQ ID NO:5)
- Ser-Gly-Lys-Pro-Ala-Ile-Ile-Pro-Asp-Arg-Glu-Val-Leu-Tyr-Arg-
Glu-Phe-Asp-Glu-Met-Glu-Glu-Cys-Ser-Gln-His-Leu-Pro-Tyr-Ile-
25 Glu-Gln-Gly-Met-Met-Leu-Ala-Glu-Gln-Phe-Lys-Gln-Lys-Ala-Leu-
Gly-Leu, (IIH; SEQ ID NO:6),
- Ser-Gly-Lys-Pro-Ala-Ile-Ile-Pro-Asp-Arg-Glu-Val-Leu-Tyr-Arg-
Glu-Phe-Asp-Glu-Met-Glu-Glu-Cys-Ser-Gln-His-Leu-Pro-Tyr-Ile,
30 (IIID; SEQ ID NO:7),

- 1 Lys-Gln-Lys-Ala-Leu-Gly-Leu-Leu-Gln-Thr-Ala-Ser-Arg-Gln-Ala-
Glu-Val-Ile-Ala-Pro-Ala-Val-Gln-Thr-Asn-Trp-Gln-Lys-Leu-Glu-
Thr-Phe-Trp-Ala-Lys-His-Met-Trp-Asn-Phe, (V; SEQ ID NO:8),
- 5 Ser-Thr-Ile-Pro-Lys-Pro-Gln-Arg-Lys-Thr-Lys-Arg-Asn-Thr-Asn-
Arg-Arg-Pro-Gln-Asp-Val-Lys-Phe Pro Gly Gly Gly Gln Ile Val-
Gly-Gly-Val-Tyr-Leu-Leu-Pro-Arg-Arg-Gly-Pro-Arg-Leu-Gly-Val-
Arg-Ala-Thr-Arg-Lys-Thr-Ser-Glu-Arg-Ser-Gln-Pro-Arg-Gly-Arg-
Arg, (VIIIIE; SEQ ID NO:9),
- 10 Asn-Asp-Arg-Val-Val-Val-Ala-Pro-Asp-Arg-Glu-Ile-Leu-Tyr-Glu-
Ala-Phe-Asp-Glu-Met-Glu-Glu-Cys-Ala-Ser-Lys-Ala-Ala-Leu-Ile-
Glu-Glu-Gly-Gln-Arg-Met-Ala-Glu-Met-Leu-Lys-Ser-Lys-Ile-Gln-
Gly-Leu, (PepA; SEQ ID NO:10),
- 15 or a sequence corresponding to one of these sequences
which is from a corresponding region in a strain or
isolate of HCV. Moreover, when the peptide moiety
comprises two or more clusters, the clusters are joined
20 by a linking group or when the clusters each have a
sequence from a different one of the above sequences,
then the clusters can be joined directly or joined by a
linking group.
- When the peptide moiety contains sequences
25 from different ones of the above sequences, such
peptides are referred to as hybrid peptides. Hybrid
peptides can but do not necessarily contain clusters.
Clusters in hybrid peptides can be joined directly or by
linking groups. In the hybrid peptides, the length of
30 contiguous amino acids from each of the sequences can be
up to about 60 residues.
- 35

1 Another aspect of the invention provides a
method of detecting antibodies to HCV or diagnosis of
HCV infection or NANBH by using an immunoeffective
amount of the subject peptide composition in an
5 immunoassay procedure, and particularly in an ELISA
procedure, or a passive hemagglutination (PHA) assay.
Immunoassays and kits for the detection and diagnosis of
NANBH and HCV infection are also provided.

 In accordance with the present invention,
10 extensive epitope analysis led to the refinement and
further definition of epitopes that are useful in the
detection and diagnosis of NANBH and HCV infection.
This analysis has established that effective diagnostic
peptides for NANBH or HCV infection are branched,
15 synthetic peptides which are hybrids of peptides
containing one or more HCV epitopes from different
peptides, also referred to herein as hybrid peptides.
Moreover, the peptides of this invention also include
branched synthetic peptides having at least one epitope
20 which is specifically immunoreactive with antibodies
against HCV and having a peptide moiety which comprises
one or more clusters of about 3 to about 20 contiguous
amino acids from the peptides designated as Pep3, Pep8,
Pep11, Pep18, Pep25, IIH, IIID, V, VIIIE, PepA, or a
25 homologous peptide from a corresponding region in
another strain or isolate of HCV. In addition, when the
peptide moiety of these peptides, also referred to
herein as cluster peptides, contain two or more
clusters, then the clusters are joined by a linking
30 group. The linking group consists of, but is not
limited to, one or more naturally occurring amino acids,
one or more unnatural amino acids, or one or more amino

1 acid analogues which can form peptidyl bonds (or
peptidyl-like bonds) and are stable to the conditions
employed during peptide synthesis. In the case of
hybrid peptides that contain clusters, the clusters can
5 be joined directly or can be joined by a linking group.

The sequences of the peptides subjected to
detailed epitope analysis, and from which the peptide
moieties of the subject branched peptides are derived,
are set forth above and are the sequences designated as
10 Pep3, Pep8, Pep11, Pep18, Pep25, IIH, IIID, V, VIIIE and
PepA or a homologous peptide from the corresponding
region in another strain or isolate of HCV, and
analogues and segments thereof.

As used herein a "cluster" is a sequence from
15 3 to about 20 contiguous amino acids from one of the
peptide sequences described herein or an analog or
segment thereof. In a preferred embodiment, a cluster
has a sequence of 3 to 9 contiguous amino acids.

The branched hybrid and cluster peptides of
20 the present invention including their analogues and
segments are useful for the detection of antibodies to
HCV in body fluids, the diagnosis of NANBH, and for the
vaccination of healthy mammals, particularly humans, to
stimulate the production of antibodies to HCV, including
25 neutralizing or protective antibodies.

The subject branched peptides can comprise
combinations or segments, i.e., longer or shorter
peptide chains by having more amino acids, including
unnatural amino acids, added to the terminal amino
30 acids, or by having amino acids removed from either
terminal end. For example, the sequence KKK (Lys-Lys-
Lys) can be added to the amino terminus of peptides.

1 Similarly, an M (methionine) residue can be placed at
the carboxy terminus of the peptide moiety, i.e. between
the peptide moiety and the branch structure.

5 As used herein "segments" means a shorter
region of a parent peptide which retains an epitope
effective in detecting NANBH-associated antibodies. For
example, C10A is a segment of VIIIE, its parent peptide.
A segment can be derived from either end of its parent
peptide or from an internal sequence of its parent
10 peptide.

The subject branched peptides can also
comprise analogues thereof to accommodate strain-to-
strain variation among different isolates of HCV or
other substitutions in the prescribed sequences which do
15 not effect immunogenicity of the epitope. HCV is
indicated to have frequent mutations. Several variant
strains/isolates are known to exist, such as PT, J, J1
and J4 [Houghton, 1989; Okamoto, 1990; Houghton, 1990;
and Kato, 1990] and it is expected that other variant
20 strains also exist. Adjustments for conservative
substitutions and selection among the alternatives where
non-conservative substitutions are involved, can be made
in the prescribed sequences. The analogues of the
branched synthetic peptides, especially the hybrid
25 peptides, can therefore comprise substitutions,
insertions and/or deletions of the recited amino acids
of the above sequence to accommodate the various
strains, as long as the immunoreactivity recognizable by
the antibodies to HCV is preserved. The substitutions
30 and insertions can be accomplished with naturally-
occurring amino acids, unnatural amino acids or amino
acid analogues capable of forming peptidyl bonds or

1 peptide-like bonds (e.g., peptide thiol analogues).
Analog peptides in accordance with this invention are
synthesized and tested against an HCV serum panel to
determine the immunoreactivity of the peptide as
5 described hereinbelow.

Further, with appropriate amino acid
modification or substitutions, it is expected that
various peptide analogues based on the prescribed amino
acid sequences can be synthesized with properties giving
10 rise to lower background readings or better binding
capacity to solid phases useful for HCV antibody
screening assays. In particular, peptides containing
unnatural amino acids can significantly reduce
background readings.

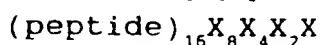
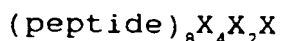
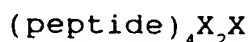
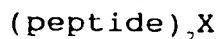
15 The subject branched peptides can also be used
to form conjugates, i.e., the peptides can be coupled
directly or indirectly, by methods known in the art, to
carrier proteins such as bovine serum albumin (BSA),
human serum albumin (HSA), or to red blood cells or
20 latex particles.

As used herein, natural amino acids are the 20
amino acids commonly found in proteins (i.e. alanine,
aspartic acid, asparagine, arginine, cysteine, glycine,
glutamine, glutamic acid, histidine, isoleucine,
25 leucine, lysine, methionine, phenylalanine, proline,
serine, threonine, tyrosine, tryptophan and valine). As
used herein the natural amino acids also include the D-
and L- forms of such amino acids.

As used herein "unnatural amino acids" include
30 both D- and L- forms of any other amino acids whether
found in a protein, whether found in nature or whether
synthetically produced. Unnatural amino acids can

1 include, but are not limited to, β -alanine, ornithine, norleucine, norvaline, hydroxyproline, thyroxine, gamma-amino butyric acid, homoserine, citrulline and the like.

The branched peptides of the present invention
5 are represented by one of the formulae:



10 wherein X is an amino acid or an amino acid analog having two amino groups and one carboxyl group, each group capable of forming a peptide bond linkage.

Preferably X is lysine or a lysine analog such as ornithine. The amino acid analog can be an α -amino
15 acid, a β -amino acid, or any other either natural or non-natural amino acid with two amino groups and one carboxyl group available for forming peptide bonds. Preferred branched peptides of the invention are dimers, tetramers and octamers, especially those having a
20 branching core structure composed of lysine, i.e. where X is lysine. Branched dimer are especially preferred.

The peptide moiety of the branched peptides can vary in length from about 10 to about 100 amino acids residues. Preferably the peptide moieties contain
25 about 17 to about 60 amino acid residues. Moreover, the hybrid and cluster peptide moieties can be optimized to the minimal overall length necessary to contain an epitope effective in detecting NANBH-associated antibodies yet still retain the superior sensitivity and
30 selectivity of the present invention.

-11-

1 The preferred branched peptides of the present
invention are provided in Table 1. The source of each
peptide is provided in Table 2.

5

10

15

20

25

30

35

1

TABLE 1
BRANCHED PEPTIDES^{a, b}

HYBRID PEPTIDES WITH OR WITHOUT CLUSTERS

5	H1	YEPFVVHGCPLPPPKSPVPPRKKRTIIPDREVLVREFDEMECSQHLPIPKPQRKTKRNTNRRPQDVKFPGGGQIVG-DIM
	H2A	LYREFDEMEDCSQHLPIPKPNRKTTRNTQRRPNDVKFPGGGNIVGM-OCT
	H2B	PDREILYREFDEMEDCSQHLPIPKPNRKTTRNTQRRPNDVKFPGGGNIVGM-OCT
	H2C	IIPDREILYREFDEMEDCSQHLPIPKPNRKTTRNTQRRPNDVKFPGGGNIVGM-OCT
	H2CK	IIPDREILYREFDEMEDCSQHLPIPKPNRKTTRNTQRRPNDVKFPGGGNIVGM-OCT
10	H2D	SGKPAIIPDREILYREFDEMEDCSQHLPIPKPNRKTTRNTQRRPNDVKFPGGGNIVGM-OCT
	H2DK	SSKPAIIPDREILYREFDEMEDCSQHLPIPKPNRKTTRNTQRRPNDVKFPGGGNIVGM-OCT
	H3	GCSGGTYDIIICDELHSTDATSIVGIGTILDQAETAGRHIFCHTKKKCDELASKLVALGM-OCT
	H4A	YEPFVVHGRHLIFCHTKKKCDELASKLVALGM-OCT
	H4B	PLVETWKKPDYEPFVVHGRHLIFCHTKKKCDELASKLVALGM-OCT
15	H6A	IEQGMMLAENFKQKALGLPRRGPRGLRATRKTTERSQPRGRM-OCT
	H6B	SGKPAIIPEREVIEQGMMLAENFKQKALGLPRRGPRGLRATRKTTERSQPRGRM-OCT
	H7	SGKPTIIPDREILYREFDEMEDCSQHLPIIDQGMMLAENFKQKALGLVKFPGGGQI-DIM
	3KH7	KKKSGKPTIIPDREILYREFDEMEDCSQHLPIIDQGMMLAENFKQKALGLVKFPGGGQI-DIM

CLUSTER PEPTIDES

20	C1A	IIPDREILYREFDEMEDCSQHLPI-DIM
	C1B	SSKPAIIPDREILYREFDEMEDCSQHLPI-DIM
	C2A	PLVETWKKPDYEPFVVH-OCT
	C2B	PLVETWKKPDYEPFVVH-OCT
	C3	KKKSGKPTIIPDREILYREFDEMEDCSQHLPIIDQGMMLAENFKQKALGL-DIM
25	C4	KKKIPKNRKTTRNTQRRPNDVKFPGGGNIVGGVYLVPRRGPRGLRATRKTTERSQPRGRM-DIM
	C5A	DCSQHLPIIDQGMMLA-DIM
	C5B	ILYREFDEMEDCSQHLPIIDQGMMLA-DIM
	C5C	SGKPTIIPDREILYREFDEMEDCSQHLPIIDQGMMLA-DIM
	3KCS5C	KKKSGKPTIIPDREILYREFDEMEDCSQHLPIIDQGMMLA-DIM
30	C6A	PLVETWKKPEYEPFVVH-DIM
	C6B	PLVETWKKPEYEPFVVH-OCT

35

-13-

1	C7A	CSQHvPYIEQGMILAEQFKQKAvGL-DIM
	C7B	LYREFDEIECSQHvPYIEQGMILAEQFKQKAvGL-DIM
	C7C	SGKPAvIPDREVLYREFDEIECSQHvPYIEQGMILAEQFKQKAvGL-DIM
	3KC7C	KKKSGKPAvIPDREVLYREFDEIECSQHvPYIEQGMILAEQFKQKAvGL-DIM
5	C8A	DYEPPVVH-DIM
	C8B	PLVETWKKDDYEPPVVH-DIM
	C8C	PLVETWKOPTYEPPVVH-DIM
	C9A	GRHLIvCHSKKKCDELAALKVALG-DIM
	C9B	EIPFYGKAvPLEvIKGGRHLIvCHSKKKCDELAALKVALG-DIM
10	C10A	RPNDvKFPGGGNIvGGVYLVPRRGPRIGLRATRKTTERSQpRGRR-DIM
	C10B	IPKPNRKTTRNTQRRPNDvKFPGGGNIvGGVYLVPRRGPRIGLRATRKTTERSQpRGRR-DIM
	3KC10B	KKKIPKPNRKTTRNTQRRPNDvKFPGGGNIvGGVYLVPRRGPRIGLRATRKTTERSQpRGRR-DIM

- a
- Abbreviations: The amino acid sequences are provided in one letter code except that unnatural amino acids are indicated by: v, norvaline; l, norleucine; p, hydroxyproline; o, ornithine. Other abbreviations are DIM, lysine dimer; OCT, lysine octamer.
- 15 b
- The branched core for these peptides is composed of lysine residues, e.g., 1 lysine for dimer peptides and 7 lysines for octamer peptides.

TABLE 2

20 SOURCE OF HYBRID AND CLUSTER BRANCHED PEPTIDES

Source Peptide	Branched Peptides from Table 1
Pep11	C2A, C2B, C6A, C6B, C8A, C8B, C8C
Pep18	C9A, C9B
IIH	C3, C5A, C5B, C5C, 3KC5C, C7A, C7B, C7C, 3KC7C
25 IIID	C1A, C1B
VIIIE	C4, C10A, C10B, 3KC10B
Pep3 + Pep18	H3
Pep11 + Pep18	H4A, H4B
Pep11 + IIID + VIIIE	H1
IIH + VIIIE	H6A, H6B, H7, 3KH7
IIID + VIIIE	H2A, H2B, H2C, H2CK, H2D, H2DK

30

35

1 The peptide compositions of the present
invention can be composed of one or more of the branched
hybrid peptides, branched cluster peptides or any
combination of such peptides. Preferably such
5 compositions contain from one to 10 branched peptides,
and even more preferably from one to four branched
peptides.

 In a preferred embodiment, the peptide
compositions of the present invention can be a mixture
10 of branched peptides (1) C3 dimer, C9B dimer, C6A dimer
and 3KH7 dimer; (2) 3K204h dimer, C4 dimer, C2B octamer;
(3) C4 dimer, C9B dimer, C6A dimer and H7 dimer; or (4)
3KH7 dimer, C6A dimer and C4 dimer. The effective ratio
of peptides for diagnosing or detecting NANBH or HCV
15 present in peptide compositions containing mixtures of
the subject peptides can be readily determined by one of
ordinary skill in the art. Typically, these ratios
range from about 1 to about 50 on a per weight basis of
peptide.

20 An especially preferred peptide composition
for diagnosis and detection of NANBH or HCV infection is
mixture (1), branched peptides 3KC10B dimer, C9B dimer,
C6A dimer and 3KH7 dimer in a weight ratio of 5:15:1:25.

 To determine the efficacy of the subject
25 peptides in detecting and diagnosing NANBH and HCV
infection, the peptides are tested for their
immunoreactivity with special specimens previously
selected through the screening of thousands of patient
and normal sera for immunoreactivity with HCV. Such
30 serum panels are commercially available and examples
thereof are provided in the Examples.

-15-

1 The strategy for serological validation
depends on the expected characteristics of the target
epitopes. For example, universal immunodominant
epitopes, such as the gp41 transmembrane peptide of HIV-
5 1, can be screened by a single representative serum
sample from a patient known to be infected with the
virus. Epitopes that are not recognized by all infected
individuals, or those for which antibody is produced
late or only transiently, and especially epitopes which
10 give rise to neutralizing antibodies, must be screened
by large panels of sera. While both methods of
screening can be employed in the present invention to
refine the epitope analysis of HCV using the subject
peptides, the latter method is particularly useful in
15 assessing the subject peptides for superior selectivity
and sensitivity.

The identification of the immunoreactive
epitopes is also dependent on the panel of sera used.
The more closely the panel represents the population
20 most likely to be seropositive for an epitope, the
greater the chance that the epitope will be identified
and thoroughly mapped. Hence, to extend the range of
reactivity of an assay comprised of previously
identified epitopes, a large number of samples from
25 individuals at risk of infection but seronegative
against known epitopes should be employed for screening.

The process of "serological validation" is
particularly difficult when the epitopes to be
identified elicit antibodies only in a subpopulation of
30 an infected patient group. When such epitopes become
targets for identification, special attention must be
paid to synthetic peptides which show very weak

1 reactivity when tested by an enzyme immunoassay or any
other immunological testing method.

In this regard, the low background absorbance
of synthetic peptides, especially peptides with
5 unnatural amino acids, allows for the precise detection
of weak reactivities. In some cases, absorbances of 50
mA versus background reading are of sufficient
significance and can lead to the identification of
important epitopes through successive refinement of the
10 amino acid sequence of a peptide. With good laboratory
practices, consistent and reliable results can be
obtained when working in the range of absorbances below
200-300 mA.

The advantages of using synthetic peptides are
15 known. Since the peptides not derived biologically from
the virus, there is no danger of exposure to a disease
causing pathogen. The peptides can be readily
synthesized using standard techniques, such as the
Merrifield method of synthesis [Merrifield (1963) J. Am.
20 Chem. Soc. 85:2149-2154]. Hence, there is no
involvement with HCV at any time during the process of
making the test reagent. Another problem which can be
minimized by using peptides rather than recombinantly
expressed proteins (or peptides) is the rate of false
25 positive results caused by the presence of antigenic
material co-purified with the HCV fusion protein. For
example, certain normal individuals have antibodies to
Escherichia coli or yeast proteins which are cross
reactive with the antigenic materials from the
30 expression system used in recombinant-based diagnostic
tests. Sera from such normal individuals show a false

1 positive reaction in such immunoassays which is
eliminated in immunoassays of the present invention.

Moreover, because the peptide compositions of
the present invention are synthetically prepared, the
5 quality can be controlled and as a result,
reproducibility of the test results can be assured.
Also, since very small amounts of a peptide are required
for each test procedure, and because the expense of
preparing a peptide is relatively low, the cost of
10 screening body fluids for antibodies to HCV and the
diagnosis of NANBH infection is relatively low.

The peptides and peptide compositions prepared
in accordance with the present invention can be used to
detect HCV infection and diagnose NANBH by using them as
15 the test reagent in an enzyme-linked immunoadsorbent
assay (ELISA), an enzyme immunodot assay, a passive
hemagglutination assay (e.g., PHA test) or other
well-known immunoassays. In accordance with the present
invention, any suitable immunoassay can be used with the
20 subject peptides. Such techniques are well known to the
ordinarily skilled artisan and have been described in
many standard immunology manuals and texts, see for
example, by Harlow et al. (1988) Antibodies: A
Laboratory Manual, Cold Spring Harbor Laboratory Press,
25 Cold Spring Harbor, NY, 726 pp. In a preferred
embodiment, the immunoassay is an ELISA using a solid
phase coated with the peptide compositions of the
present invention. ELISA techniques are well known in
the art. In another preferred embodiment the
30 immunoassay is a PHA assay.

The immunoassays of the present invention are
used to screen body fluids and tissues for the presence

1 of NANBH or HCV and thereby to detect such agents and
aid the practitioner in diagnosis of NANBH or HCV
infection. The body fluids which can be subjected to
such screening include blood and blood fractions (e.g.
5 serum), saliva, or any other fluid which contains
antibodies against HCV.

Another aspect of the present invention is
directed to a kit for the detection and diagnosis of
NANBH or HCV infection in mammalian body fluids (e.g.
10 serum, tissue extracts, tissue fluids), in vitro cell
culture supernatants, and cell lysates. The kit can be
compartmentalized to receive a first container adapted
to contain one or more of the peptides (i.e. a peptide
composition) of this invention.

15 Preferably the kit of this invention is an
ELISA or a PHA test kit for detection or diagnosis of
NANBH or HCV infection. For an ELISA test kit, the kit
contains (a) a container (e.g., a 96-well plate) having
a solid phase coated with one of the subject peptide
20 compositions; (b) a negative control sample; (c) a
positive control sample; (d) specimen diluent and (e)
antibodies to human IgG, which antibodies are labelled
with a reporter molecule. If the reporter molecule is
an enzyme, then the kit also contains a substrate for
25 said enzyme.

In an exemplified use of the subject kit, a
sample to be tested is contacted with a mammalian body
fluid, diluted in sample diluent if necessary, for a
time and under conditions for any antibodies, if
30 present, to bind to the peptide contained in the
container. After removal of unbound material (e.g. by
washing with sterile phosphate buffered saline), the

1 secondary complex is contacted with labelled antibodies
to human IgG. These antibodies bind to the secondary
complex to form a tertiary complex and, since the second
antibodies are labeled with a reporter molecule, when
5 subjected to a detecting means, the tertiary complex is
detected. The reporter molecule can be an enzyme,
radioisotope, fluorophore, bioluminescent molecule,
chemiluminescent molecule, biotin, avidin, streptavidin
or the like. For ELISA the reporter is preferably an
10 enzyme.

The examples serve to illustrate the present
invention and are not to be used to limit the scope of
the invention.

15 EXAMPLE 1

Detection of antibodies
to the core region of HCV in early
seroconversion sample using branched cluster peptides

The wells of 96-well plates were coated
separately for 1 hour at 37° with 1 µg/ml of peptide
20 using 100 µL per well in 10mM NaHCO₃ buffer, pH 9.5, for
each of two branched peptides from the core region of
HCV (peptide C4, Table 1; and test peptide T1 related to
VIIIE and having the sequence
KKKIPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSEERS
25 QPRGRR-DIM

The peptide-coated wells were then incubated
with 250 µL of 3% by weight of gelatin in PBS in 37°C
for 1 hour to block non-specific protein binding sites,
followed by three washes with PBS containing 0.05% by
30 volume of TWEEN 20 and then dried. The test specimens
containing HCV antibody positive patient sera were

1 diluted with PBS containing 20% by volume normal goat
serum, 1% by weight gelatin and 0.05% by volume TWEEN 20
at dilutions of 1:20 volume to volume, respectively.
200 μ L of the diluted specimens were added to each of
5 the wells and allowed to react for 15 minutes at 37°C.

The wells were then washed six times with
0.05% by volume TWEEN 20 in PBS in order to remove
unbound antibodies. Horseradish peroxidase conjugated
goat anti-human IgG was used as a second antibody tracer
10 to bind with the HCV antibody-peptide antigen complex
formed in positive wells. 100 μ L of peroxidase labeled
goat anti-human IgG at a dilution of 1:1800 in 1% by
volume normal goat serum, 0.05% by volume TWEEN 20 in
PBS was added to each well and incubated at 37°C for
15 another 15 minutes.

The wells were washed six times with 0.05% by
volume TWEEN 20 PBS to remove unbound antibody and
reacted with 100 μ L of the substrate mixture containing
0.04% by weight orthophenylenediamine (OPD) and 0.12% by
20 volume hydrogen peroxide in sodium citrate buffer, pH
5.0. This substrate mixture was used to detect the
peroxidase label by forming a colored product.
Reactions were stopped by the addition of 100 μ L of 1.0M
 H_2SO_4 and the A_{492nm} measured.

25 The sensitivity of these two peptides in
detecting antibody to the core region was tested with a
seroconversion panel in which the earliest antibody
response is known to be against core (Serologicals Panel
4813, Donor 02190D as referenced in U.S. Patent No.
30 5,106,726; early core response as referenced in Hosein,
1991). The bleed date chosen for comparison was August
30, 1988. The optical density obtained with peptide C4

-21-

1 was 0.320 and with T1, 0.512. Both peptides were more
sensitive than the linear peptide VIIIE with three
lysine residues at its N terminus when coated at the
same concentration, in which case the absorbance on the
5 same sample was 0.075.

EXAMPLE 2

Branched hybrid peptides confer improved sensitivity and
specificity relative to the individual peptides

10 The immunoreactivity of branched hybrid
peptide 3KH7 (Table 1) containing an epitope from the
NS-4 and core regions of HCV was tested on panel 3
containing 41 known NANBH samples using the ELISA assay
format as described in Example 1. Table 3 shows that
15 this hybrid peptide retained the reactivity of both the
NS-4 and the core regions as compared to octamer T2
(related to VIIIE) from the core region only and peptide
T3 (SEQ ID NO:11; related to IIH) from the NS-4 region
only. Furthermore, sample 3-35 showed improved
20 reactivity with the hybrid peptide relative to either
single region peptide.

The specificity of the hybrid peptide 3KH7 was
tested on a panel of 48 random blood donor samples
screened negative for antibodies to HCV. Only one of
25 the negative samples had an absorbance greater than
0.200 A with the hybrid peptide, whereas twenty percent
of these samples had absorbance values greater than
0.200 A with the octamer T2. Branched cluster peptide
C3, containing an epitope from the NS-4 region but
30 lacking the core epitopes, gave absorbance values
greater than 0.200 A on 5/48 negative samples.
Therefore the combination of epitopes from the two

- 1 regions as presented in the hybrid peptide resulted in improved specificity for detection of NANBH.

Table 3

5

HCV Positive Sample ^a	A _{492nm} (pos/neg)		
	3KH7	T2 ^b	T3 ^b
3-2	0.491(+)	0.068(-)	0.756(+)
3-10	1.164(+)	0.027(-)	1.857(+)
3-21	2.576(+)	0.095(-)	2.226(+)
10 3-32	1.653(+)	1.188(+)	2.236(+)
3-35	2.303(+)	0.800(+)	0.324(+)
3-39	1.441(+)	0.486(+)	1.676(+)
3-7	1.118(+)	3.229(+)	0.582(+)
3-8	0.696(+)	1.860(+)	0.003(-)
3-9	1.408(+)	2.797(+)	0.163(-)
3-12	1.870(+)	0.328(+)	0.037(-)
15 3-26	1.607(+)	3.233(+)	0.355(+)

^a The remaining samples in panel 3 were negative on all peptides or showed no improvement in using the branched hybrid peptide compared with the test peptides.

20 ^b The sequences of control peptides T2 and T3 are, respectively, VKFPGGGQIM-octamer and

KKKSGKPAIIPDREVLVREFDEMEEC SQHLPYIEQGMMLAEQFKQKALGL.

25

EXAMPLE 3

Comparison of sensitivity and specificity in detection of NANBH-associated antibodies in branched cluster peptides with unnatural amino acids linking groups

30 The immunoreactivity of branched cluster peptide C10B (Table 1) from the core region with clusters separated by unnatural amino acids was compared

35

1 with a similar peptide T1 (Example 1) lacking such
unnatural amino acids, using panel 3 samples in an ELISA
assay format as described in Example 1. Table 4
illustrates seven samples in which the absorbance for
5 the peptide containing unnatural amino acids was higher
than for the corresponding peptide lacking unnatural
amino acids, i.e., branched peptide C10B was more
sensitive than T1. The specificity of these two
peptides was equivalent with 0/48 negative samples
10 having absorbance readings greater than 0.200 A.

The immunoreactivity of branched cluster
peptide C8C (Table 1) from the NS-5 region of HCV having
clusters separated by unnatural amino acids was compared
with the corresponding branched peptide lacking
15 unnatural amino acids (C6A dimer; this peptide has
clusters separated by natural amino acids; Table 1).
Both peptides detected 18/41 samples from panel 3 as
positive. Table 5 shows six samples in which the
absorbance with the peptide containing unnatural amino
20 acids was higher than for the corresponding peptide
lacking unnatural amino acids.

Table 6 shows four reactive samples from panel
3 in which peptide 3KC7C (Table 1) had increased
absorbance values compared to peptide C3 (Table 1),
25 i.e., the presence of unnatural amino acids imparted
greater sensitivity to the assay for detection of NANBH
and HCV.

Furthermore, a marked improvement in
specificity, measured by the ELISA procedure as
30 described in Example 1, was also obtained with branched
cluster peptide 3KC7C from the NS-4 region of HCV having
clusters separated by unnatural amino acids. With

- 1 peptide 3KC7C, 0/48 negative samples had absorbance
values greater than 0.200 A, whereas 5/48 had absorbance
values greater than 0.200 A with branched peptide C3
which lacked unnatural amino acids but had natural amino
5 acid separating the clusters. Specificity was also
improved by addition of the unnatural amino acid in
peptide C8C, in that only 1/48 negative random donor
samples had absorbance readings greater than 0.200 A,
compared with 2/48 for peptide C6A.

10

Table 4

HCV Positive Sample ^a	A _{492nm}	
	C10B	T1
15 3-7	2.451	2.005
3-8	1.081	0.873
3-9	2.665	2.272
3-12	0.446	0.352
3-24	2.378	2.088
3-25	2.399	1.555
20 3-39	1.289	0.767

^a See Table 6

Table 5

HCV Positive Sample ^a	A _{492nm}	
	C8C	C6A
25 3-1	1.622	1.246
3-5	2.130	1.907
3-11	0.895	0.782
3-27	2.710	2.463
3-33	2.108	1.763
30 3-36	2.236	2.016

^a See Table 6

35

1

Table 6

HCV Positive Sample ^a	A _{492nm}	
	3KC7C	C3
3-7	0.389	0.350
3-14	2.034	1.670
3-29	1.561	1.350
3-41	> 3.0	2.570

5

10

^a For Tables 4-6, the remaining samples in panel 3 were negative on both peptides or showed no improvement in using the branched hybrid peptide compared to the test or control peptides.

EXAMPLE 4

Improved NS-5 immunoreactivity conferred by a shorter branched peptide relative to its linear parent peptide

15

A 17 residue branched octamer cluster peptide, C2A from the NS-5 region of HCV (Table 1), was able to detect antibody in all 23/41 samples from panel 3 that were reactive with its parent linear peptide T4, a 44 residue peptide having the sequence
ARPDYNPPLVETWKKPDYYYEPPVHGCPLPPPKSPPVPPPRKKRT SEQ ID NO:12). Table 7 shows five samples from panel 3 that exhibited higher absorbance values with peptide octamer C2A than with linear peptide T4.

20

25

30

35

1

Table 7

5

HCV Positive Sample ^a	A _{492nm}	
	T4	C2A
3-7	0.742	1.377
3-11	1.188	1.815
3-16	3.139	3.745
3-26	2.263	2.527
3-33	2.118	2.631

10

^a The remaining samples in panel 3 were negative on both peptides or showed no improvement in using 17-mer compared with the 44-mer.

15

EXAMPLE 5

Earlier detection of NANBH-associated antibodies in a seroconversion panel using a mixture of branched peptides

A mixture of dimer peptides, 3KC10B, 3KH7, C9B and C6A (1, 5, 3, 0.25 ug/ml, respectively) was coated on wells of 96-well plates and assayed using the ELISA procedure described in Example 1. The sequence of each branched peptide is provided in Table 1. The sensitivity of this mixture was compared with that of Format C peptides (described in EPO 0468527 A2 and consisting of peptides IIH, V and VIIIE coated at 5, 3 and 2 μ g/ml, respectively) using seroconversion panel 4813 described in Example 1. Table 8 shows that seroconversion samples were consistently positive on the mixture of peptides one week before antibody was detected by Format C. Earlier samples at bleed dates of August 9 and August 16, 1988 show fluctuation of

35

-27-

- 1 antibody response near the cutoff of the assay and
indicate detection of passive antibodies from the
transfusion of this patient that occurred July 19, 1988.

5

Table 8

Panel	Donor	Bleed Date	ALT ^a (u/L)	EIA Ratio		
				Format C	Mixture ^b	
10	1	02190D	880809	40.0	0.108	1.197
			880816	32.0	0.045	0.899
			880823	32.0	0.025	1.044
			880830	180.0	1.037	1.197
			880928	401.0	7.193	3.303
			881109	NA	10.185	10.250
			881122	NA	9.770	11.548

- 15 ^a Abbreviations: ALT = Alanine amino-transferase
^b The composition of Format C and Mixture are described in
Example 5

20

25

30

35

-28-

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: United Biomedical Inc.
- (ii) TITLE OF INVENTION: Novel Branched Hybrid and Cluster Peptides
Effective in Diagnosing and Detecting Non-A,
Non-B Hepatitis
- (iii) NUMBER OF SEQUENCES: 12
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: UNITED BIOMEDICAL INC.
 - (B) STREET: 25 Davids Drive
 - (C) CITY: Hauppauge
 - (D) STATE: New York
 - (E) COUNTRY: USA
 - (F) ZIP: 11788
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: M. Lisa Wilson
 - (B) REGISTRATION NUMBER: 34,045
 - (C) REFERENCE/DOCKET NUMBER: 9055
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 516-273-2828
 - (B) TELEFAX: 516-273-1717
 - (C) TELEX:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Leu His
1 5 10 15
Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln
20 25 30
Ala Glu Thr Ala Gly
35

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Phe Thr Phe Ser Pro Arg Arg His Trp Thr Thr Gln Gly Cys Asn Cys
 1 5 10 15
 Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met
 20 25 30
 Met Met Asn Trp Ser Pro Thr Ala
 35 40

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

-30-

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Glu	Ile	Leu	Arg	Lys	Ser	Arg	Arg	Phe	Ala	Gln	Ala	Leu	Pro	Val	Trp
1				5					10					15	
Ala	Arg	Pro	Asp	Tyr	Asn	Pro	Pro	Leu	Val	Glu	Thr	Trp	Lys	Lys	Pro
			20					25					30		
Asp	Tyr	Glu	Pro	Pro	Val	Val	His	Gly	Cys	Pro	Leu	Pro	Pro	Pro	Lys
		35					40					45			
Ser	Pro	Pro	Val	Pro	Pro	Pro	Arg	Lys	Lys	Arg	Thr				
	50					55					60				

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Glu	Ile	Pro	Phe	Tyr	Gly	Lys	Ala	Ile	Pro	Leu	Glu	Val	Ile	Lys	Gly
1				5				10						15	
Gly	Arg	His	Leu	Ile	Phe	Cys	His	Ser	Lys	Lys	Lys	Cys	Asp	Glu	Leu
			20					25					30		
Ala	Ala	Lys	Leu	Val	Ala	Leu									
			35												

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

-31-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Pro Val Val Pro Gln Ser Phe Gln Val Ala His Leu His Ala Pro Thr
 1 5 10 15
 Gly Ser Gly Lys Ser
 20

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu
 1 5 10 15
 Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu Pro Tyr Ile Glu Gln
 20 25 30
 Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys Ala Leu Gly Leu
 35 40 45

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu
 1 5 10 15
 Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu Pro Tyr Ile
 20 25 30

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- 32 -

- (A) LENGTH: 40 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Lys Gln Lys Ala Leu Gly Leu Leu Gln Thr Ala Ser Arg Gln Ala Glu
1 5 10 15

Val Ile Ala Pro Ala Val Gln Thr Asn Trp Gln Lys Leu Glu Thr Phe
20 25 30

Trp Ala Lys His Met Trp Asn Phe
35 40

(2) INFORMATION FOR SEQ ID NO:9:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 61 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Ser Thr Ile Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg
1 5 10 15

Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly
20 25 30

Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala Thr
35 40 45

Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg
50 55 60

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

-33-

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Asn	Asp	Arg	Val	Val	Val	Ala	Pro	Asp	Arg	Glu	Ile	Leu	Tyr	Glu	Ala
1				5					10					15	
Phe	Asp	Glu	Met	Glu	Glu	Cys	Ala	Ser	Lys	Ala	Ala	Leu	Ile	Glu	Glu
			20					25					30		
Gly	Gln	Arg	Met	Ala	Glu	Met	Leu	Lys	Ser	Lys	Ile	Gln	Gly	Leu	
		35					40					45			

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Lys	Lys	Lys	Ser	Gly	Lys	Pro	Ala	Ile	Ile	Pro	Asp	Arg	Glu	Val	Leu
1				5					10					15	
Tyr	Arg	Glu	Phe	Asp	Glu	Met	Glu	Glu	Cys	Ser	Gln	His	Leu	Pro	Tyr
			20					25					30		
Ile	Glu	Gln	Gly	Met	Met	Leu	Ala	Glu	Gln	Phe	Lys	Gln	Lys	Ala	Leu
		35					40					45			
Gly	Leu														
	50														

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

-34-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Ala	Arg	Pro	Asp	Tyr	Asn	Pro	Pro	Leu	Val	Glu	Thr	Trp	Lys	Lys	Pro
1				5					10					15	
Asp	Tyr	Tyr	Tyr	Glu	Pro	Pro	Val	Val	His	Gly	Cys	Pro	Leu	Pro	Pro
			20					25					30		
Pro	Lys	Ser	Pro	Pro	Val	Pro	Pro	Pro	Arg	Lys	Lys	Arg	Thr		
		35					40					45			

1 WE CLAIM:

1. A peptide composition comprising at least one branched peptide represented by the formula

(peptide)₂X

5 (peptide)₄X₂X

(peptide)₈X₄X₂X

(peptide)₁₆X₈X₄X₂X

wherein X is an amino acid or an amino acid analog having two amino groups and one carboxyl group, each group capable of forming a peptide bond linkage, and said peptide moiety comprises at least one epitope which is specifically immunoreactive with NANBH-associated antibodies, wherein said peptide moiety comprises at least one cluster of from about 3 to about 15 20 contiguous amino acids selected from the group of sequences consisting of SEQ ID NOS: 1 to 10 (Pep3, Pep8, Pep11, Pep18, Pep25, IIH, IIID, V, VIIIE, PepA) and a sequence corresponding to one of said sequences which is from a corresponding region in a strain or isolate of HCV;

20 when said peptide moiety comprises two or more clusters, said clusters are joined by a linking group, said linking group being at least one natural amino acid, unnatural amino acid, or amino acid analog, and when said two or more clusters have sequences from a 25 different one of the above sequences, said clusters can be joined directly or can be joined by said linking group; and

further wherein said peptide moiety comprises about 10 to about 100 amino acids.

30 2. The peptide composition of Claim 1 comprising a mixture of two or more of said peptides.

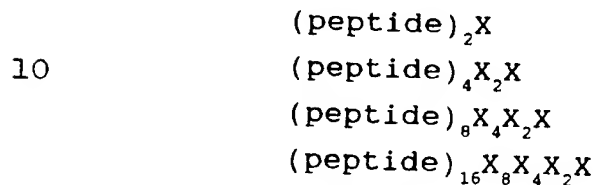
-36-

- 1 3. The peptide composition of Claim 1,
 wherein said peptide is conjugated to a carrier.
4. The peptide composition of Claim 1,
 wherein said cluster comprises from 5 to 9 contiguous
5 amino acids.
5. The peptide composition of Claim 1
 wherein said peptide moiety further comprises a segment
 of one of said sequences.
6. The peptide composition of Claim 1
10 wherein said sequence is the sequence designated as SEQ
 ID NO:3 (Pep11).
7. The peptide composition of Claim 6
 wherein said peptide is C2A, C2B, C6A, C6B, C8A, C8B or
 C8C.
- 15 8. The peptide composition of Claim 1
 wherein said sequence is the sequence designated as SEQ
 ID NO:4 (Pep18).
9. The peptide composition of Claim 8
 wherein said peptide is C9A or C9B.
- 20 10. The peptide composition of Claim 1
 wherein said sequence is the sequence designated as SEQ
 ID NO:6 (IIH).
11. The peptide composition of Claim 10
 wherein said peptide is C3, C5A, C5B, C5C, 3KC5C, C7A,
25 C7B, C7C or 3KC7C.
12. The peptide composition of Claim 1
 wherein said sequence is the sequence designated as SEQ
 ID NO:7 (IIID).
- 30 13. The peptide composition of Claim 12
 wherein said peptide is C1A or C1B.

1 14. The peptide composition of Claim 1
 wherein said sequence is the sequence designated as SEQ
 ID NO:9 (VIIIIE).

5 15. The peptide composition of Claim 14
 wherein said peptide is C4, C10A, C10B, or 3KC10B.

16. A peptide composition comprising at
 least one branched hybrid peptide represented by the
 formula



10 wherein X is an amino acid or an amino acid analog
 having two amino groups and one carboxyl group, each
 15 group capable of forming a peptide bond linkage, and
 said peptide moiety comprises at least one
 epitope which is specifically immunoreactive with
 antibodies against HCV, wherein said peptide moiety
 comprises a first sequence from one of the following
 20 sequences and one or more additional sequences, each
 from a different one of said sequences, wherein said
 sequence is selected from the group of sequences
 consisting of SEQ ID NOS: 1 to 10 (Pep3, Pep8, Pep11,
 Pep18, Pep25, IIH, IIID, V, VIIIIE, PepA), a sequence
 25 corresponding to one of said sequences which is from a
 corresponding region in a strain or isolate of HCV, an
 analog of one of said sequences, and a segment of one of
 said sequences; and

 further wherein said peptide moiety comprises
 30 about 10 to about 100 amino acids.

 17. The peptide composition of Claim 16
 wherein said sequences are the sequences designated as

1 SEQ ID NO:3, SEQ ID NO:7 and SEQ ID NO:9 (Pep11, IIID and VIIIE).

18. The peptide composition of Claim 17 wherein said peptide is H1.

5 19. The peptide composition of Claim 16 wherein said sequences are the sequences designated as SEQ ID NO:1 and SEQ ID NO:4 (Pep3 and Pep18).

20. The peptide composition of Claim 19 wherein said peptide is H3.

10 21. The peptide composition of Claim 16 wherein said sequences are the sequences designated as SEQ ID NO:3 and SEQ ID NO:4 (Pep11 and Pep18).

22. The composition of Claim 21 wherein said peptide is H4A or H4B.

15 23. The peptide composition of Claim 16 wherein said sequences are the sequences designated as SEQ ID NO:6 and SEQ ID NO:9 (IIH and VIIIE).

24. The peptide composition of Claim 23 wherein said peptide is H6A, H6B, H7 or 3KH7.

20 25. The peptide composition of Claim 16 wherein said sequences are the sequences designated as SEQ ID NO:7 and SEQ ID NO:9 (IIID and VIIIE).

25 26. The peptide composition of Claim 25 wherein said peptide is H2A, H2B, H2C, H2CK, H2D or H2DK.

27. A peptide composition comprising peptides 3KC10B, C9B, C6A and 3KH7.

28. A peptide composition comprising peptides 3KH7, C6A and C4.

30 29. A peptide composition comprising peptides C3, C4 and C2B.

1 30. A peptide composition comprising
peptides C4, C9B, C6A and 3KH7.

31. A peptide of any one of Claims 1 to 30.

5 32. A method of detecting NANBH-associated
antibodies which comprises using an effective amount of
a peptide composition according to any one of Claims 1
to 30 in an immunoassay procedure.

10 33. A method of detecting NANBH or HCV
infection which comprises contacting an effective amount
of a peptide composition of any one of Claims 1 to 30
with a body fluid, tissue or tissue extract in an
immunoassay procedure for a time sufficient to form a
complex between said peptide composition and any
antibody in said fluid, said tissue, or said tissue
15 extract, and subjecting said complex to a detecting
means.

34. The method of Claim 32 or 33 wherein
said immunoassay procedure is an ELISA or a PHA
procedure.

20 35. A kit for detection or diagnosis of
NANBH or HCV infection comprising a first container
adapted to contain the peptide composition of any one of
Claims 1 to 30.

25 36. The kit of Claim 35 wherein said kit is
an ELISA or PHA test kit.

37. An ELISA test kit for detection and
diagnosis of NANBH or HCV infection comprising

30 (a) a container having a solid phase coated
with the peptide composition of any one of Claims 1 to
30;

(b) a negative control sample;

(c) a positive control sample;

-40-

- 1 (d) specimen diluent; and
 (e) antibodies to human IgG, said antibodies
 labeled with a reporter molecule.

5

10

15

20

25

30

35

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/08638**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(5) :C07K 7/00; A61K 39/12; C12Q 1/70

US CL :530/324; 424/89; 435/5

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/324; 424/89; 435/5; 436/820

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, Dialog

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 5,106,726 (Wang) 21 April 1992, entire document, especially col. 11, lines 10-27; col. 28, lines 50-66; and Example 7.	1-33, 35-37
Y	EP, A, 0,318,216 (Houghton et al.) 31 May 1989, see entire document.	1-33, 35-37

☐ Further documents are listed in the continuation of Box C.
 ☐ See patent family annex.

* Special categories of cited documents:	*T	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be part of particular relevance	*X*	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y*	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Z*	document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means		
P document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

14 December 1993

Date of mailing of the international search report

JAN 03 1994

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Authorized officer

DONNA C. WORTMAN, PH.D.

Facsimile No. NOT APPLICABLE

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No
PCT/US93/08638**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons.

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 34
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : C12N 15/12, A61K 37/02 C07K 13/00, C12N 1/21 // (C12N 1/21, C12R 1:19)	A1	(11) International Publication Number: WO 94/00571 (43) International Publication Date: 6 January 1994 (06.01.94)
(21) International Application Number: PCT/GB93/01282 (22) International Filing Date: 16 June 1993 (16.06.93) (30) Priority data: 9213376.8 24 June 1992 (24.06.92) GB 9304057.4 1 March 1993 (01.03.93) GB (71) Applicant (for all designated States except US): SMITH-KLINE BEECHAM PLC [GB/GB]; New Horizons Court, Brentford, Middlesex TW8 9EP (GB). (72) Inventors; and (75) Inventors/Applicants (for US only) : SMITH, Richard, Anthony, Godwin [GB/GB]; DODD, Ian [GB/GB]; FREEMAN, Anne, Mary [GB/GB]; MOSSAKOWSKA, Danuta, Eva, Irena [GB/GB]; SmithKline Beecham Pharmaceuticals, Great Burgh, Yew Tree Bottom Road, Epsom, Surrey KT18 5XQ (GB).		(74) Agent: VALENTINE, Jill, Barbara; SmithKline Beecham, Corporate Patents, Great Burgh, Yew Tree Bottom Road, Epsom, Surrey KT18 5XQ (GB). (81) Designated States: JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>
(54) Title: SOLUBLE CRI DERIVATIVES		
(57) Abstract A soluble polypeptide comprising, in sequence, one to four short consensus repeats (SCR) selected from SCR 1,2,3 and 4 of long homologous repeat A(LHR-A) as the only structurally and functionally intact SCR domains of CRI and including at least SCR3.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NE	Niger
BE	Belgium	GN	Guinea	NL	Netherlands
BF	Burkina Faso	GR	Greece	NO	Norway
BG	Bulgaria	HU	Hungary	NZ	New Zealand
BJ	Benin	IE	Ireland	PL	Poland
BR	Brazil	IT	Italy	PT	Portugal
BY	Belarus	JP	Japan	RO	Romania
CA	Canada	KP	Democratic People's Republic of Korea	RU	Russian Federation
CF	Central African Republic	KR	Republic of Korea	SD	Sudan
CG	Congo	KZ	Kazakhstan	SE	Sweden
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovak Republic
CM	Cameroon	LU	Luxembourg	SN	Senegal
CN	China	LV	Latvia	TD	Chad
CS	Czechoslovakia	MC	Monaco	TG	Togo
CZ	Czech Republic	MG	Madagascar	UA	Ukraine
DE	Germany	ML	Mali	US	United States of America
DK	Denmark	MN	Mongolia	UZ	Uzbekistan
ES	Spain			VN	Viet Nam
FI	Finland				

SOLUBLE CR1 DERIVATIVES.

The present invention relates to polypeptides and their use in the diagnosis and therapy of disorders involving complement activity and various inflammatory and immune disorders.

Constituting about 10% of the globulins in normal serum, the complement system is composed of many different proteins that are important in the immune system's response to foreign antigens. The complement system becomes activated when its primary components are cleaved and the products alone or with other proteins, activate additional complement proteins resulting in a proteolytic cascade. Activation of the complement system leads to a variety of responses including increased vascular permeability, chemotaxis of phagocytic cells, activation of inflammatory cells, opsonization of foreign particles, direct killing of cells and tissue damage. Activation of the complement system may be triggered by antigen-antibody complexes (the classical pathway) or, for example, by lipopolysaccharides present in cell walls of pathogenic bacteria (the alternative pathway).

Complement receptor type 1 (CR1) has been shown to be present on the membranes of erythrocytes, monocytes/macrophages, granulocytes, B cells, some T cells, splenic follicular dendritic cells, and glomerular podocytes. CR1 binds to the complement components C3b and C4b and has also been referred to as the C3b/C4b receptor. The structural organisation and primary sequence of one allotype of CR1 is known (Klickstein *et al.*, 1987, J. Exp. Med. 165:1095-1112, Klickstein *et al.*, 1988, J. Exp. Med. 168:1699-1717; Hourcade *et al.*, 1988, J. Exp. Med. 168:1255-1270, WO 89/09220, WO 91/05047). It is composed of 30 short consensus repeats (SCRs) that each contain around 60-70 amino acids. In each SCR, around 29 of the average 65 amino acids are conserved. Each SCR has been proposed to form a three dimensional triple loop structure through disulphide linkages with the third and first and the fourth and second half-cystines in disulphide bonds. CR1 is further arranged as 4 long homologous repeats (LHRs) of 7 SCR each. Following a leader sequence, the CR1 molecule consists of the N-terminal LHR-A, the next two repeats, LHR-B and LHR-C, and the most C-terminal LHR-D followed by 2 additional SCRs, a 25 residue putative transmembrane region and a 43 residue cytoplasmic tail.

Based on the mature CR1 molecule having a predicted N-terminal glutamine residue, hereinafter designated as residue 1, the first four SCR domains of LHR-A are defined herein as consisting of residues 2-58, 63-120, 125-191 and 197-252, respectively, of mature CR1.

Hourcade *et al.*, 1988, J. Exp. Med. 168:1255-1270 observed an alternative polyadenylation site in the human CR1 transcriptional unit that was predicted to

produce a secreted form of CR1. The mRNA encoded by this truncated sequence comprises the first 8.5 SCRs of CR1, and encodes a protein of about 80 kDa which was proposed to include the C4b binding domain. When a cDNA corresponding to this truncated sequence was transfected into COS cells and expressed, it demonstrated the expected C4b binding activity but did not bind to C3b (Krych *et al.*, 1989, FASEB J. 3:A368; Krych *et al.* Proc. Nat. Acad. Sci. 1991, 88, 4353-7). Krych *et al.*, also observed a mRNA similar to the predicted one in several human cell lines and postulated that such a truncated soluble form of CR1 with C4b binding activity may be synthesised in humans.

10 In addition, Makrides *et al.* (1992, J. Biol. Chem. 267 (34) 24754-61) have expressed SCR 1 + 2 and 1 + 2 + 3 + 4 of LHR-A as membrane-attached proteins in CHO cells.

Several soluble fragments of CR1 have also been generated via recombinant DNA procedures by eliminating the transmembrane region from the DNAs being expressed (WO 89/09220, WO 91/05047). The soluble CR1 fragments were functionally active, bound C3b and/or C4b and demonstrated Factor I cofactor activity depending upon the regions they contained. Such constructs inhibited *in vitro* complement-related functions such as neutrophil oxidative burst, complement mediated hemolysis, and C3a and C5a production. A particular soluble construct, sCR1/pBSCR1c, also demonstrated *in vivo* activity in a reversed passive Arthus reaction (WO 89/09220, WO 91/05047; Yeh *et al.*, 1991, J. Immunol. 146:250), suppressed post-ischemic myocardial inflammation and necrosis (WO 89/09220, WO 91/05047; Weisman *et al.*, Science, 1990, 249:146-1511; Dupe, R. *et al.* Thrombosis & Haemostasis (1991) 65(6) 695.) and extended survival rates following transplantation (Pruitt & Bollinger, 1991, J. Surg. Res 50:350; Pruitt *et al.*, 1991 Transplantation 52; 868). Furthermore, co-formulation of sCR1/pBSCR1c with p-anisoylated human plasminogen-streptokinase-activator complex (APSAC) resulted in similar anti-haemolytic activity as sCR1 alone, indicating that the combination of the complement inhibitor sCR1 with a thrombolytic agent was feasible (WO 91/05047).

Soluble polypeptides corresponding to part of CR1 have now been found to possess functional complement inhibitory, including anti-haemolytic, activity.

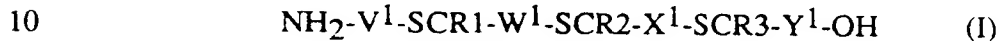
According to the present invention there is provided a soluble polypeptide comprising, in sequence, one to four short consensus repeats (SCR) selected from SCR 1, 2, 3 and 4 of long homologous repeat A (LHR-A) as the only structurally and functionally intact SCR domains of CR1 and including at least SCR3.

In preferred aspects, the polypeptide comprises, in sequence, SCR 1, 2, 3 and 4 of LHR-A or SCR 1, 2 and 3 of LHR-A as the only structurally and functionally

intact SCR domains of CR1.

It is to be understood that variations in the amino acid sequence of the polypeptide of the invention by way of addition, deletion or conservative substitution of residues, including allelic variations, in which the biological activity of the polypeptide is retained, are encompassed by the invention. Conservative substitution is understood to mean the retention of the charge and size characteristics of the amino acid side chain, for example arginine replaced by histidine.

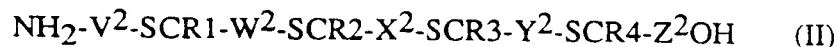
In one aspect, the polypeptide of the invention may be represented symbolically as follows:



in which SCR1 represents residues 2-58 of mature CR1, SCR2 represents residues 63-120 of mature CR1, SCR3 represents residues 125-191 of mature CR1, and V^1 , W^1 , X^1 and Y^1 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

In a preferred embodiment of formula (I), W^1 , X^1 and Y^1 represent residues 59-62, 121-124 and 192-196, respectively, of mature CR1 and V^1 represents residue 1 of mature CR1 optionally linked via its N-terminus to methionine.

In another aspect the polypeptide of the invention may be represented symbolically as follows:



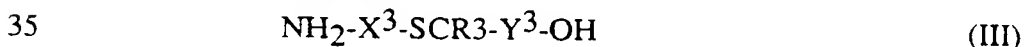
in which SCR1, SCR2 and SCR3 are as hereinbefore defined, SCR4 represents residues 197-252 of mature CR1 and V^2 , W^2 , X^2 , Y^2 and Z^2 represents bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

In preferred embodiments of formula (II), W^2 , X^2 , Y^2 and Z^2 represent residues 59-62, 121-124, 192-196, and residues 253 respectively, of mature CR1 and V^2 represents residue 1 of mature CR1 optionally linked via its N-terminus to methionine.

In one particular embodiment of formula (II) arginine 235 is replaced by histidine.

In the preferred embodiment of formula (II), residue 235 is arginine.

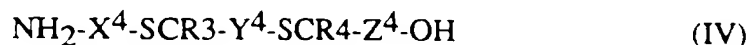
In one further aspect, the polypeptide of the invention may be represented symbolically as follows:



in which SCR3 is as hereinbefore defined and X^3 and Y^3 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

In a preferred embodiment of formula (III) X^3 represents amino acids 122-124 of mature CR1 optionally linked to methionine at its N-terminus and Y^4 represents amino acids 192-196 of mature CR1.

5 In another further aspect, the polypeptide of the invention may be represented symbolically as follows:



in which SCR3 and SCR4 are as hereinbefore defined and X^4 , Y^4 and Z^4 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

10 In a preferred embodiment of formula (IV) X^4 represents amino acids 122-124 of mature CR1 optionally linked to methionine at its N-terminus and Y^4 and Z^4 represent amino acids 192-196 and 253 respectively of mature CR1.

In a further aspect, the invention provides a process for preparing a CR1 polypeptide according to the invention which process comprises expressing DNA
15 encoding said polypeptide in a recombinant host cell and recovering the product.

In particular, the process may comprise the steps of:

- i) preparing a replicable expression vector capable, in a host cell, of expressing a DNA polymer comprising a nucleotide sequence that encodes said polypeptide;
- ii) transforming a host cell with said vector;
- 20 iii) culturing said transformed host cell under conditions permitting expression of said DNA polymer to produce said polypeptide; and
- iv) recovering said polypeptide.

The DNA polymer comprising a nucleotide sequence that encodes the polypeptide also forms part of the invention.

25 The process of the invention may be performed by conventional recombinant techniques such as described in Sambrook *et al.*, Molecular Cloning : A laboratory manual 2nd Edition. Cold Spring Harbor Laboratory Press (1989) and DNA Cloning vols I, II and III (D. M. Glover ed., IRL Press Ltd).

30 The invention also provides a process for preparing the DNA polymer by the condensation of appropriate mono-, di- or oligomeric nucleotide units.

The preparation may be carried out chemically, enzymatically, or by a combination of the two methods, *in vitro* or *in vivo* as appropriate. Thus, the DNA polymer may be prepared by the enzymatic ligation of appropriate DNA fragments, by conventional methods such as those described by D. M. Roberts *et al.*, in
35 Biochemistry 1985, 24, 5090-5098.

The DNA fragments may be obtained by digestion of DNA containing the required sequences of nucleotides with appropriate restriction enzymes, by chemical synthesis, by enzymatic polymerisation, or by a combination of these methods.

Digestion with restriction enzymes may be performed in an appropriate buffer at a temperature of 20°-70°C, generally in a volume of 50µl or less with 0.1-10µg DNA.

Enzymatic polymerisation of DNA may be carried out *in vitro* using a DNA polymerase such as DNA polymerase I (Klenow fragment) in an appropriate buffer containing the nucleoside triphosphates dATP, dCTP, dGTP and dTTP as required at a temperature of 10°-37°C, generally in a volume of 50µl or less.

Enzymatic ligation of DNA fragments may be carried out using a DNA ligase such as T4 DNA ligase in an appropriate buffer at a temperature of 4°C to 37°C, generally in a volume of 50µl or less.

The chemical synthesis of the DNA polymer or fragments may be carried out by conventional phosphotriester, phosphite or phosphoramidite chemistry, using solid phase techniques such as those described in 'Chemical and Enzymatic Synthesis of Gene Fragments - A Laboratory Manual' (ed. H.G. Gassen and A. Lang), Verlag Chemie, Weinheim (1982), or in other scientific publications, for example M.J.Gait, H.W.D. Matthes M. Singh, B.S. Sproat and R.C. Titmas, Nucleic Acids Research, 1982, 10, 6243; B.S. Sproat and W. Bannwarth, Tetrahedron Letters, 1983, 24, 5771; M.D. Matteucci and M.H. Caruthers, Tetrahedron Letters, 1980, 21, 719; M.D. Matteucci and M.H. Caruthers, Journal of the American Chemical Society, 1981, 103, 3185; S.P. Adams *et al.*, Journal of the American Chemical Society, 1983, 105, 661; N.D. Sinha, J. Biernat, J. McMannus and H. Koester, Nucleic Acids Research, 1984, 12, 4539; and H.W.D. Matthes *et al.*, EMBO Journal, 1984, 3, 801. Preferably an automated DNA synthesiser (for example, Applied Biosystems 381A Synthesiser) is employed.

The DNA polymer is preferably prepared by ligating two or more DNA molecules which together comprise a DNA sequence encoding the polypeptide.

The DNA molecules may be obtained by the digestion with suitable restriction enzymes of vectors carrying the required coding sequences.

The precise structure of the DNA molecules and the way in which they are obtained depends upon the structure of the desired product. The design of a suitable strategy for the construction of the DNA molecule coding for the polypeptide is a routine matter for the skilled worker in the art.

In particular, consideration may be given to the codon usage of the particular host cell. The codons may be optimised for high level expression in *E. coli* using the principles set out in Devereux *et al.*, (1984) Nucl. Acid Res., 12, 387.

The expression of the DNA polymer encoding the polypeptide in a recombinant host cell may be carried out by means of a replicable expression vector capable, in the host cell, of expressing the DNA polymer. The expression vector is

novel and also forms part of the invention.

The replicable expression vector may be prepared in accordance with the invention, by cleaving a vector compatible with the host cell to provide a linear DNA segment having an intact replicon, and combining said linear segment with one or
5 more DNA molecules which, together with said linear segment, encode the polypeptide, under ligating conditions.

The ligation of the linear segment and more than one DNA molecule may be carried out simultaneously or sequentially as desired.

Thus, the DNA polymer may be preformed or formed during the construction
10 of the vector, as desired. The choice of vector will be determined in part by the host cell, which may be prokaryotic, such as *E. coli*, or eukaryotic, such as mouse C127, mouse myeloma, chinese hamster ovary, fungi e.g. filamentous fungi or unicellular 'yeast' or an insect cell such as *Drosophila*. The host cell may also be in a transgenic animal. Suitable vectors include plasmids, bacteriophages, cosmids and recombinant
15 viruses derived from, for example, baculoviruses or vaccinia.

The DNA polymer may be assembled into vectors designed for isolation of stable transformed mammalian cell lines expressing the fragment e.g. bovine papillomavirus vectors in mouse C127 cells, or amplified vectors in chinese hamster ovary cells (DNA Cloning Vol. II D.M. Glover ed. IRL Press 1985; Kaufman, R.J. *et al.*
20 *et al.* Molecular and Cellular Biology 5, 1750-1759, 1985; Pavlakis G.N. and Hamer, D.H. Proceedings of the National Academy of Sciences (USA) 80, 397-401, 1983; Goeddel, D.V. *et al.*, European Patent Application No. 0093619, 1983).

The preparation of the replicable expression vector may be carried out conventionally with appropriate enzymes for restriction, polymerisation and ligation
25 of the DNA, by procedures described in, for example, Sambrook *et al.* cited above. Polymerisation and ligation may be performed as described above for the preparation of the DNA polymer. Digestion with restriction enzymes may be performed in an appropriate buffer at a temperature of 20°-70°C, generally in a volume of 50µl or less with 0.1-10µg DNA.

30 The recombinant host cell is prepared, in accordance with the invention, by transforming a host cell with a replicable expression vector of the invention under transforming conditions. Suitable transforming conditions are conventional and are described in, for example, Sambrook *et al.*, cited above, or "DNA Cloning" Vol. II, D.M. Glover ed., IRL Press Ltd, 1985.

35 The choice of transforming conditions is determined by the host cell. Thus, a bacterial host such as *E. coli*, may be treated with a solution of CaCl₂ (Cohen *et al.*, Proc. Nat. Acad. Sci., 1973, 69, 2110) or with a solution comprising a mixture of RbCl, MnCl₂, potassium acetate and glycerol, and then with 3-[N-morpholino]-

propane-sulphonic acid, RbCl and glycerol or by electroporation as for example described by Bio-Rad Laboratories, Richmond, California, USA, manufacturers of an electroporator. Mammalian cells in culture may be transformed by calcium co-precipitation of the vector DNA onto the cells or by using cationic liposomes.

5 The invention also extends to a host cell transformed with a replicable expression vector of the invention.

Culturing the transformed host cell under conditions permitting expression of the DNA polymer is carried out conventionally, as described in, for example, Sambrook *et al.*, and "DNA Cloning" cited above. Thus, preferably the host cell is supplied with nutrient and cultured at a temperature below 45°C.

10 The protein product is recovered by conventional methods according to the host cell. Thus, where the host cell is bacterial such as *E. coli* and the protein is expressed intracellularly, it may be lysed physically, chemically or enzymatically and the protein product isolated from the resulting lysate. Where the host cell is
15 mammalian, the product is usually isolated from the nutrient medium.

Where the host cell is bacterial, such as *E. coli*, the product obtained from the culture may require folding for optimum functional activity. This is most likely if the protein is expressed as inclusion bodies. There are a number of aspects of the isolation and folding process that are regarded as important. In particular, the
20 polypeptide is preferably partially purified before folding, in order to minimise formation of aggregates with contaminating proteins and minimise misfolding of the polypeptide. Thus, the removal of contaminating *E. coli* proteins by specifically isolating the inclusion bodies and the subsequent additional purification prior to folding are important aspects of the procedure.

25 The folding process is carried out in such a way as to minimise formation of intermediate-folded states of the polypeptide. Thus, careful consideration needs to be given to, among others, the salt type and concentration, temperature, protein concentration, redox buffer concentrations and duration of folding. The exact condition for any given polypeptide generally cannot be predicted and must be
30 determined by experiment.

There are numerous methods available for the folding of proteins from inclusion bodies and these are known to the skilled worker in this field. The methods generally involve breaking all the disulphide bonds in the inclusion body, for example with 50mM 2-mercaptoethanol, in the presence of a high concentration of denaturant
35 such as 8M urea or 6M guanidine hydrochloride. The next step is to remove these agents to allow folding of the proteins to occur. Formation of the disulphide bridges requires an oxidising environment and this may be provided in a number of ways, for example by air, or by incorporating a suitable redox system, for example a mixture of

reduced and oxidised glutathione.

Preferably, the inclusion body is solubilised using 8M urea, in the presence of mercaptoethanol, and protein is folded, after initial removal of contaminating proteins, by addition of cold buffer. A preferred buffer is 20mM ethanolamine

5 containing 1mM reduced glutathione and 0.5mM oxidised glutathione. The folding is preferably carried out at a temperature in the range 1 to 50°C over a period of 1 to 4 days.

If any precipitation or aggregation is observed, the aggregated protein can be removed in a number of ways, for example by centrifugation or by treatment with
10 precipitants such as ammonium sulphate. Where either of these procedures are adopted, monomeric polypeptide is the major soluble product.

If the bacterial cell secretes the protein, folding is not usually necessary.

The polypeptide of this invention is useful in the treatment or diagnosis of many complement-mediated or complement-related diseases and disorders including,
15 but not limited to, those listed below.

Disease and Disorders Involving Complement

Neurological Disorders

multiple sclerosis
20 stroke
Guillain Barré Syndrome
traumatic brain injury
Parkinson's disease
allergic encephalitis
25 Alzheimer's disease

Disorders of Inappropriate or Undesirable Complement Activation

haemodialysis complications
hyperacute allograft rejection
30 xenograft rejection
corneal graft rejection
interleukin-2 induced toxicity during IL-2 therapy
paroxysmal nocturnal haemoglobinuria

35 Inflammatory Disorders

inflammation of autoimmune diseases
Crohn's Disease
adult respiratory distress syndrome

- thermal injury including burns or frostbite
- uveitis
- psoriasis
- asthma
- 5 acute pancreatitis

Post-Ischemic Reperfusion Conditions

- myocardial infarction
- balloon angioplasty
- 10 atherosclerosis (cholesterol-induced) & restenosis
- hypertension
- post-pump syndrome in cardiopulmonary bypass or renal haemodialysis
- renal ischemia
- intestinal ischaemia

15

Infectious Diseases or Sepsis

- multiple organ failure
- septic shock

20 Immune Complex Disorders and Autoimmune Diseases

- rheumatoid arthritis
- systemic lupus erythematosus (SLE)
- SLE nephritis
- proliferative nephritis
- 25 glomerulonephritis
- haemolytic anemia
- myasthenia gravis

Reproductive Disorders

- 30 antibody- or complement-mediated infertility

Wound Healing

- 35 The present invention is also directed to a pharmaceutical composition comprising a therapeutically effective amount of a polypeptide, as above, and a pharmaceutically acceptable carrier or excipient.

The present invention also provides a method of treating a disease or disorder associated with inflammation or inappropriate complement activation comprising

administering to a subject in need of such treatment a therapeutically effective amount of a polypeptide of this invention.

In the above methods, the subject is preferably a human.

An effective amount of the polypeptide for the treatment of a disease or
5 disorder is in the dose range of 0.01-100mg/kg; preferably 0.1mg-10mg/kg.

For administration, the polypeptide should be formulated into an appropriate pharmaceutical or therapeutic composition. Such a composition typically contains a therapeutically active amount of the polypeptide and a pharmaceutically acceptable excipient or carrier such as saline, buffered saline, dextrose, or water. Compositions
10 may also comprise specific stabilising agents such as sugars, including mannose and mannitol, and local anaesthetics for injectable compositions, including, for example, lidocaine.

Further provided is the use of a polypeptide of this invention in the manufacture of a medicament for the treatment of a disease or disorder associated
15 with inflammation or inappropriate complement activation.

In order to inhibit complement activation and, at the same time, provide thrombolytic therapy, the present invention provides compositions which further comprise a therapeutically active amount of a thrombolytic agent. An effective amount of a thrombolytic agent is in the dose range of 0.01-10mg/kg; preferably 0.1-
20 5mg/kg. Preferred thrombolytic agents include, but are not limited to, streptokinase, human tissue type plasminogen activator and urokinase molecules and derivatives, fragments or conjugates thereof. The thrombolytic agents may comprise one or more chains that may be fused or reversibly linked to other agents to form hybrid molecules (EP-A-0297882 and EP 155387), such as, for example, urokinase linked to plasmin
25 (EP-A-0152736), a fibrinolytic enzyme linked to a water-soluble polymer (EP-A-0183503). The thrombolytic agents may also comprise muteins of plasminogen activators (EP-A-0207589). In a preferred embodiment, the thrombolytic agent may comprise a reversibly blocked *in vitro* fibrinolytic enzyme as described in U.S. Patent No. 4,285,932. A most preferred enzyme is a p-anisoyl plasminogen-streptokinase
30 activator complex as described in U.S. Patent No. 4,808,405, and marketed by SmithKline Beecham Pharmaceuticals under the Trademark EMINASE (generic name anistreplase, also referred to as APSAC; Monk *et al.*, 1987, Drugs 34:25-49).

Routes of administration for the individual or combined therapeutic compositions of the present invention include standard routes, such as, for example,
35 intravenous infusion or bolus injection. Active complement inhibitors and thrombolytic agents may be administered together or sequentially, in any order.

The present invention also provides a method for treating a thrombotic condition, in particular acute myocardial infarction, in a human or non-human animal.

This method comprises administering to a human or animal in need of this treatment an effective amount of a polypeptide according to this invention and an effective amount of a thrombolytic agent.

Also provided is the use of a polypeptide of this invention and a thrombolytic agent in the manufacture of a medicament for the treatment of a thrombotic condition in a human or animal. Such methods and uses may be carried out as described in WO 91/05047.

This invention further provides a method for treating adult respiratory distress syndrome (ARDS) in a human or non-human animal. This method comprises administering to the patient an effective amount of a polypeptide according to this invention.

The invention also provides a method of delaying hyperacute allograft or hyperacute xenograft rejection in a human or non-human animal which receives a transplant by administering an effective amount of a polypeptide according to this invention. Such administration may be to the patient or by application to the transplant prior to implantation.

The invention yet further provides a method of treating wounds in a human or non-human animal by administering by either topical or parenteral e.g. intravenous routes, an effective amount of a polypeptide according to this invention.

GENERAL METHODS USED IN EXAMPLES

(i) DNA cleavage

Cleavage of DNA by restriction endonucleases was carried out according to the manufacturer's instructions using supplied buffers. Double digests were carried out simultaneously if the buffer conditions were suitable for both enzymes. Otherwise double digests were carried out sequentially where enzyme requiring the lowest salt concentration was added first to the digest. Once that digest was complete the salt concentration was altered and the second enzyme added.

(ii) Production of blunt ended DNA fragments

The recessed 3' termini of DNA fragments were filled in using the Klenow fragment of DNA polymerase I as described in Sambrook *et al* (1989).

(iii) DNA purification/ concentration and analysis

Removal of protein contaminants, nucleosides was with phenol/CHCl₃ followed by precipitation with ethanol. DNA was analysed on horizontal agarose gel electrophoresis: both methods are described in Sambrook *et al* (1989).

(iv) DNA fragment isolation

1. DNA purification on DEAE NA45 membranes

DNA fragments were purified from agarose gels by making an incision in the

agarose above and just below the required DNA fragment. NA45 membranes from Schleicher & Schuell (Anderman, Great Britain) that had been soaked in TE (10 mM Tris pH 8.0, 1 mM EDTA) were inserted into the incisions and current reapplied to the gel until the DNA fragment was trapped on the lower membrane; higher
5 molecular weight DNA was trapped on the upper membrane. The lower membrane was removed from the gel and the DNA eluted into 0.05 M arginine/1 M NaCl at 70°C for 2 hours. The DNA was then concentrated by ethanol precipitation as described in Sambrook *et al* (1989).

2. Electroelution

10 DNA fragments were excised from agarose gels and DNA extracted by electroelution using the Unidirectional Electroeluter (IBI Ltd., Cambridge, England) according to the manufacturer's instructions.

3. Gel purification

15 DNA fragments were excised from agarose gels and DNA extracted using the QIAEX gel extraction kit according to the manufacturers instructions (QIAGEN Inc., USA).

(v) Plasmid preparation

Large scale plasmid preparation of plasmid DNA was carried out using CsCl as described in Sambrook *et al* (1989) or using Magic Maxipreps (Promega Corporation, Madison, USA) according to the manufacturers instructions. Mini-
20 plasmid preparations were carried out using either the alkaline lysis method described in Sambrook *et al* (1989) or Magic Minipreps (Promega Corporation, Madison, USA) according to the manufacturer's instructions.

(vi) Introduction of plasmid DNA into *E. coli*

25 1. Plasmids were transformed into *E. coli* HB101 or *E. coli* BL21 (DE3) (Studier and Moffat, 1986) that had been made competent using calcium chloride as described in Sambrook *et al* (1989).

2. Alternatively plasmids were introduced into *E. coli* DH1 (Low, 1968) or *E. coli* BL21 (DE3) by electroporation using the Gene Pulsar and Pulse Controller of Bio-
30 Rad (Bio-Rad Laboratories, Richmond, California, USA) according to the manufacturer's instructions.

(vii) Kinasing of oligonucleotides

Oligonucleotides or annealed oligonucleotides possessing 5' overhangs were kinased using T₄ polynucleotide kinase as described in Sambrook *et al* (1989).

35 (viii) Annealing and ligation of oligonucleotides

Oligonucleotides were annealed together by mixing generally equimolar concentrations of the complementary oligonucleotides in 10 mM Tris pH 8.5, 5 mM MgCl₂ and placing at 100°C for 5 minutes and then cooling very slowly to room

temperature. Annealed oligonucleotides with sticky ends were ligated to vector or other oligonucleotides containing complementary sticky ends using T₄ DNA ligase as described in Sambrook *et al* (1989).

(ix) **PCR (Polymerase Chain Reaction) amplification of DNA**

- 5 DNA fragments from ligation reactions or DNA fragments excised and purified from agarose gels were amplified by PCR from two primers complementary to the 5' ends of the DNA fragment. Approximately 0.1 - 1 µg of ligation reaction or the purified DNA from the agarose gel was mixed in 10 mM Tris pH 8.3 (at 25°C), 50 mM KCl, 0.1% gelatin; MgCl₂ concentrations were varied from 1.5 mM to 6 mM
- 10 to find a suitable concentration for each reaction. Both primers were added to a final concentration of 2 µM; each dNTP was added to a final concentration of 0.2 mM. The final reaction volume was either 75 µl or 100 µl, which was overlaid with mineral oil to prevent evaporation. Thermal cycling was then started on a thermal cycler eg. Hybaid Thermal reactor, and a typical example of conditions used was
- 15 94°C 7 mins, 45°C 2 mins, hold at 45°C for less than 5 min., and then add 5 units of Taq DNA polymerase (purchased from a commercial source, e.g. Gibco). The DNA fragment was amplified by cycling the temperature at 72°C 2 mins, 94°C 1 min and 45°C 2 min a total of 35 times.

(x) **DNA sequencing using the double stranded method**

- 20 Sequencing was carried out using "Sequenase™" (United States Biochemical Corporation) essentially as described in the manufacturer's instructions.

(xi) **DNA sequence analysis and manipulation**

Analysis of sequences were carried out on a digital VAX computer using the GCG package of programmes as described in Devereux *et al* (1984).

25 (xii) **Production of oligonucleotides**

1. Oligonucleotides were synthesised using a Gene Assembler Plus (Pharmacia LKB Biotechnology, Milton Keynes, England) or a 381A Synthesiser (Applied BioSystems) according to the manufacturer's instructions.
 2. Oligonucleotide purification was carried out either using MonoQ as
- 30 recommended by Pharmacia or by UV shadowing where recovery of synthetic oligonucleotides was by electrophoresis through a denaturing polyacrylamide gel. The oligonucleotides were loaded onto a 12% acrylamide/7M urea gel and run at 1500V until the oligonucleotide had migrated approximately two thirds of the length of the gel. The DNA was visualised using a hand-held, long-wavelength ultraviolet
- 35 lamp; and the DNA bands excised. The oligonucleotide was recovered using Sep-Pak C18 reverse phase columns (Waters) as described in Sambrook *et al* (1989).

(xiii) **Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE)**

SDS PAGE was carried out generally using the Novex system (British Biotechnology) according to the manufacturer's instructions. Prepacked gels of acrylamide concentrations 14%, 16%, 4 - 20% or 10 - 27% were the ones most frequently used. Samples for electrophoresis, including protein molecular weight standards (LMW Kit, Pharmacia) were usually diluted in 1%(w/v)SDS - containing buffer (with or without 5%(v/v) 2-mercaptoethanol), and left at room temperature for about 0.5 to 1 h before application to the gel.

(xiv) **Alteration of codon usage**

The non random use of synonymous codons has been demonstrated in *E. coli* and there is some evidence to support the belief that protein production from genes containing non-optimal or minor codons (particularly at the 5' end of the gene) is less efficient than that from genes with no such codons (e.g. Chen and Inouye, 1990). A codon usage table compiled from genes highly expressed in *E. coli* (supplied as part of the GCG sequence analysis software package, Devereux *et al*, (1984)) was used to determine the optimal codons for expression in *E. coli*. All of the first 30 codons of all constructs (where compatible with restriction enzyme sites) were optimised for high level expression. The codons for the seven amino acids: arg, gly, ile, leu, pro, ser, ala were optimised (where compatible with restriction enzyme sites) throughout the coding sequence.

(xv) **Construction of vector pBROC413**

The plasmid pT7-7 (Tabor, 1990) contains DNA corresponding to nucleotides 2065-4362 of pBR322 and like pBR322 can be mobilized by a conjugative plasmid in the presence of a third plasmid ColK. A mobility protein encoded by ColK acts on the *nic* site at nucleotide 2254 of pBR322 initiating mobilization from this point. pT7-7 was digested with *LspI* and *BglII* and the protruding 5' ends filled in with the Klenow fragment of DNA Polymerase I. The plasmid DNA fragment was purified by agarose gel electrophoresis, the blunt ends ligated together and transformed into *E. coli* DH1 by electroporation. The resultant plasmid pBROC413 (Fig.1) was identified by restriction enzyme analysis of plasmid DNA.

The deletion in pBROC413 from the *LspI* site immediately upstream of the $\phi 10$ promoter to the *BglII* site at nucleotide 434 of pT7-7 deletes the DNA corresponding to nucleotides 2065-2297 of pBR322. The *nic* site and adjacent sequences are therefore deleted making pBROC413 non mobilizable.

(xvi) **Haemolytic assay**

The anti-haemolytic activity of polypeptides was assessed by measuring the inhibition of complement mediated lysis of sheep erythrocytes sensitised with rabbit antibodies (obtained from Diamedix Corporation, Miami, USA). Human serum diluted 1:125 in 0.1 M Hepes/ 0.15 M NaCl pH 7.4 buffer was the source of

complement and was prepared from a pool of volunteers essentially as described in (Dacie & Lewis, 1975). Briefly, blood was warmed to 37°C for 5 minutes, the clot removed and the remaining serum clarified by centrifugation. The serum fraction was split into small aliquots and stored at -196°C. Aliquots were thawed as required and diluted in the Hepes buffer immediately before use.

Inhibition of complement-mediated lysis of sensitised sheep erythrocytes was measured using a standard haemolytic assay using a v-bottom microtitre plate format as follows.

50 µl of a range of concentrations (0.01-100 µg/ml but typically 0.05 - 25 µg/ml) of test protein diluted in Hepes buffer were incubated with 50 µl of the diluted serum for 15 minutes at 37°C. 100 µl of prewarmed sensitised sheep erythrocytes were added for 1 hour at 37°C in a final reaction volume of 200 µl. Samples were spun at 300g at 4°C for 15 minutes before transferring 150 µl of supernatant to flat bottomed microtitre plates and determining the absorption at 410 nm, which reflects the amount of lysis in each test solution. Maximum lysis was determined by incubating serum with erythrocytes in the absence of any inhibitor from which the proportion of background lysis had been subtracted (determined by incubating erythrocytes with buffer). The background lysis by inhibitor was assessed by incubating inhibitor with erythrocytes and then subtracting that from test samples. Inhibition was expressed as a fraction of the total cell lysis such that IH50 represents the concentration of inhibitor required to give 50% inhibition of lysis.

(xvii) **C3a RIA Assay**

Activation of complement pathways can be followed by measuring the release of the anaphylatoxin, C3a and its breakdown product C3a des Arg. Both products can be measured using a competitive radio-immuno assay purchased from Amersham International plc, U. K., (human complement C3a des Arg [¹²⁵I]assay, code RPA 518).

(a) **Alternative pathway activation by Zymosan A**

The alternative pathway of complement was activated with zymosan A, a complex carbohydrate from yeast (Sigma, catalogue number Z-4250). Zymosan A was made 50 mg/ml in Hepes buffer (0.1M Hepes/0.15M NaCl pH 7.4) or in PBS (50 mM sodium phosphate/0.1 M NaCl pH 7.4) and vortexed until a fine suspension had formed. Serum (prepared as described for the haemolytic assay; Method xvi) was preincubated with different concentrations of complement inhibitor diluted in Hepes buffer for 15 mins at 37°C using the volumes given below. Zymosan A was then vortexed for a few seconds each time before addition to the samples after which samples were incubated for a further 30 mins at 37°C. The zymosan A was then spun down at approximately 11,000g for 30 seconds at ambient temperature.

Typically 100 µl of supernatant were added to an equal volume of precipitating solution provided in the kit and the subsequent supernatant assayed as described in the technical bulletin supplied by Amersham with the C3a des Arg assay RIA kit. Each sample was assayed in duplicate and useful dilutions of the supernatant, to ensure that sample counts were on the standard curve, were found to be 1/50 - 1/100. EDTA or Futhan were not used in any solutions or tubes as suggested in the technical bulletin.

Each sample was counted for 1 minute on an LKB-Wallac 1272 Clinigamma. Data was processed using the RiaCalc program for RIA assays as supplied with the Clinigamma. The data was computed essentially as described in the Amersham technical bulletin with the standrd curve constructed by non-linear regression fit to the data.

The miniaturised assay was carried out essentially as described above but using smaller total volumes for the activation of serum.

Volumes of samples added

	serum	inhibitor	Zymosan A
Normal Assay	79 µl	20 µl	21 µl
Miniaturised Assay	26.3 µl	6.7 µl	7 µl

In the minaturised assay, after activation, typically 25 µl of the sample were precipitated. The assay kit reagent additions were reduced from 50 µl to 10 µl which enabled the assay to be carried out in a U-bottom microtitre plate containing separate detachable wells. The assay was then carried out as described in the technical bulletin using the adjusted volumes until the last dilution in isotonic saline. In this instance 200 µl of saline were added and the plate spun at approximately 2500g for 12 mins at 4°C. The supernatants from each well were carefully removed by aspiration and the precipitate was washed with a further 300 µl of isotonic saline. The plate was then spun again at about 2500g for 5 mins, 4°C and the supernatant was discarded. Wells were then counted for 10 mins each on the Clinigamma. The data was processed as above.

To determine the % inhibition of maximum activation at each inhibitor concentration, a number of controls were carried out with each experiment. These included maximum activation (A) *i.e.* serum + zymosan A only, background activation (B) *i.e.* serum + buffer only, and background activation in the presence of inhibitor (C) *i.e.* serum + inhibitor only. The background activation was generally subtracted from the maximum activation. Similarly the background activation in the presence of inhibitor was subtracted from the value of activated serum in the presence of inhibitor. These values could then be used to determine the % inhibition at each

inhibitor concentration, using the following formula:

$$\frac{1 - (D - C)}{(A - B)} \times 100$$

where D is the value of activation of serum in the presence of inhibitor and zymosan

- 5 A. The IC₅₀ is defined as the concentration of inhibitor required to reduce maximum activation by 50%.

(b) **Classical pathway activation by heat aggregated IgG**

Activation of the classical pathway by IgG was performed as follows.

- Human γ -globulin (Sigma, catalogue number G-4386) was made 14 mg/ml in 0.1 M
 10 Hepes/0.15 M NaCl pH 7.4 and heated at 60°C for 1 hour. Samples of heat aggregated IgG were then stored as small aliquots at -80°C until required. Serum was activated using heat aggregated IgG using the same volumes as described for the zymosan A normal or miniaturised assay. Preincubation of inhibitor with serum was for 15 mins at 37°C followed by addition of the heat aggregated IgG. Incubation
 15 was continued for a further 45 mins at 37°C. The samples were then assayed directly for C3a levels using either the normal or miniaturised assay.

(xviii) **C5a RIA Assay**

- Activation of complement pathways can be followed by measuring the release of the anaphylatoxin C5a and its breakdown product C5a des Arg. Both products can
 20 be measured using a competitive radio-immuno assay purchased from Amersham International plc, U. K., (human complement C5a des Arg [¹²⁵I]assay, code RPA 520).

- The alternative pathway of complement was activated with zymosan A, as described for the C3a RIA assay (Method (xvii)). The assay was carried out in the
 25 miniaturised form as described for the C3a assay using the reagents provided in the C5a des-Arg RIA kit.

- References in the Examples to amino acid numbering relate to the corresponding residues of mature CR1 protein.

30

Example 1 Construction of plasmid pDB1010-D11 encoding SCR 1 + 2

General points

- A DNA sequence for SCR 1 + 2 corresponding to amino acid 1 and ending at amino acid 124 of mature human complement receptor 1 was designed such that the
 35 5' end of the gene contained an *Nde*I site. This site comprises an ATG codon to give the initiating methionine required for the start of mRNA translation and places the gene an optimum distance from the Shine-Dalgarno ribosome binding sequence of pBROC413. The 3' end of the gene finished on two stop codons followed by a

*Hind*III site.

Restriction endonucleases that do not cut pBROC413 and that were commercially available were identified. The sequences of the restriction sites recognised by the endonucleases were translated into all three reading frames. The sites that contained rarely used codons for *E. coli* expression were discarded. The remaining sites were matched with the DNA coding for SCR 1 + 2. If the restriction site could be fitted into the DNA sequence so as to preserve the coding sequence and not add a rarely used codon, the DNA sequence was altered to include this restriction site. 10 unique restriction sites were so identified and incorporated. To enable intracellular expression of protein in *E. coli*, an ATG codon was added to the 5' end of the gene immediately preceding the codon for the first amino acid of mature CR-1. The codon ATG is part of the *Nde*I restriction site which can be used for cloning into vectors such as pBROC413. The codon corresponding to proline 124 of mature CR-1 has been changed to one encoding glutamine, which also encompasses an *Eco*RI site.

(a) **Construction of plasmid**

Oligonucleotides coding for SCR 1 + 2 (Table 1; 1 - 8) were synthesised as 4 complementary pairs of 87 - 101 mers that could be ligated in a unique fashion via complementary 8 bp overhangs between the pairs of oligonucleotides. The four complementary pairs of oligonucleotides were designated Pair A (oligos 1+2), Pair B (oligos 3+4), Pair C (oligos 5+6) and Pair D (oligos 7+8). Pair A which corresponded to the 5' end of the gene contained an *Nde*I restriction site overhang and Pair D contained a *Hind*III restriction site overhang at the 3' end. All oligonucleotides apart from 1 and 2 were purified on Pharmacia Mono Q columns prior to use. Oligonucleotide 2 of pair A and oligonucleotide 7 of pair D were kinased before annealing with their unkinased complementary oligonucleotides 1 and 8 respectively. Oligonucleotides pairs B and C were annealed first and then kinased. The kinased oligonucleotide pairs were ligated Pair A (approx. 0.1µg) to Pair B (approx. 0.2µg) and Pair C (approx. 2µg) to Pair D (approx. 4µg). The ligated oligonucleotides (A+B) were in turn ligated to (C+D) to form the gene coding for SCR 1 + 2.

The DNA coding for SCR 1 + 2 was amplified by PCR using two oligonucleotides (Table 1; 15 and 16) complementary to the two strands of DNA. Both oligonucleotides contained 5' unmatched ends that contained 6 bp of random sequence followed by the sequence of either *Nde*I or *Hind*III restriction sites followed by 18 bp complementary to the gene. Following PCR, a band of approximately 400 bp was visualised on horizontal agarose gel electrophoresis, which was excised and purified on DEAE NA45 membranes. The DNA was then cut with *Nde*I and *Hind*III

before ligating into pBROC413 that had been cut with the same enzymes. The vector was transformed into *E. coli* HB101 made competent with calcium chloride. Mini-plasmid preparations were made and the plasmid DNA was analysed by digestion with *NdeI* and *HindIII*. Plasmids containing the correct sized insert, were further subjected to restriction mapping with *EcoRI*, *HpaI*, *KpnI* and *SmaI*. The plasmids that displayed the correct restriction maps were analysed by DNA sequencing of both strands across the gene coding for SCR 1 + 2. Plasmid pDB1010-D11 was identified as having the correct sequence across the gene coding for SCR 1 + 2.

Example 2 Construction, expression, purification, folding and formulation of MQ1 -> K196 of CR-1 (SCR 1+2+3)

General Points

The DNA coding for SCR 1+2+3 was constructed by ligating DNA coding for SCR 1+2 (Example 1a) to DNA encoding SCR 3.

General points relating to SCR 3 are presented in Examp.

The SCR 3 coding unit corresponding to amino acid 122 and ending at amino acid 196 of mature CR1, was designed such that 5' end of the unit contained the *EcoRI* site at the junction of SCR's 2 & 3 as well as an *NdeI* site 5' to the *EcoRI* site. The 3' end of the unit finished on two stop codons followed by a *HindIII* site. The plasmids containing the SCR 3 coding unit and the SCR 1+2 coding unit were digested with *EcoRI* and *HindIII*. The SCR 3 coding unit was isolated and inserted downstream of the SCR 1+2 coding unit in the *EcoRI/HindIII*-cut SCR 1+2 - containing plasmid, to give a plasmid containing the SCR 1+2+3 coding unit, which corresponds to amino acids 1 to 196 of mature CR1. The addition of the SCR 3 coding unit through the *EcoRI* site, converts the codon corresponding to a glutamine at position 124 back to the authentic amino acid (proline) that is found in CR1.

(a) Construction of plasmid pDB1013-5-4 encoding SCR 1 + 2 + 3

Three pairs of oligonucleotides (Table 1; 9 - 14) encompassing the SCR coding sequence were synthesised. The oligonucleotides were first annealed (9, 10; 11,12; 13,14) and the middle pair kinased thus allowing the three pairs to be ligated together via 8 base pair overlapping sequences. The 5' end of this molecule was designed to be complementary to *NdeI* digested DNA and the 3' end to *HindIII* digested DNA. This enabled the trimer to be cloned into *NdeI/HindIII* digested pBROC413 generating pBROC435 (Fig.2). The identity of pBROC435 was checked by restriction enzyme analysis and confirmed by DNA sequencing.

Plasmid DNA from pBROC435 and pDB1010-D11 (Example 1) were both cut with *EcoRI* and *HindIII*; the *EcoRI/HindIII* band of pBROC435 coding for SCR 3

was purified on an DEAE NA45 membrane as was the cut vector pDB1010-D11. The SCR 3 coding unit was then ligated into pDB1010-D11 to generate pDB1013-5 which was then transformed into calcium chloride competent *E. coli* HB101. The resulting colonies were analysed by mini-plasmid preparation of DNA followed by restriction mapping. One of the colonies, termed pDB1013-5-4 (Fig.2), contained the SCR 1+2+3 coding unit. This plasmid was then analysed for expression of the gene product.

(b) **Expression of SCR 1 + 2 + 3**

pDB1013-5-4 was transformed into calcium chloride competent *E. coli* BL21(DE3) and resulting colonies analysed by restriction digestion of mini-plasmid DNA preparations. Single colonies were inoculated into universals containing 10 ml of L broth or NCYZM medium and 50 µg/ml ampicillin and allowed to grow overnight at 37°C, 220 r.p.m. The overnight cultures (typically 5ml) were used to inoculate each of 2 L conical flasks containing 500 mls of NCYZM medium, 150 µg/ml ampicillin; cultures were grown at 37°C, 220 r.p.m. until A₆₀₀ was 0.5 absorbance units. Cultures were induced with 1 mM isopropylthioβ-D-galactoside (IPTG) and allowed to grow a further 3 hours under the same conditions. The cultures were centrifuged (approx. 8000g/10 min) and the supernatants discarded. The cell pellets were stored at -40°C. L broth was 1% (w/v) Bactotryptone, 0.5% (w/v) Bactoyeastextract, 0.5% (w/v) NaCl. NCYZM media was L-broth containing 0.1% (w/v) casamino acids and 0.2%(w/v) MgSO₄.7H₂O, pH 7.0.

(c) **Isolation of solubilised inclusion bodies**

Frozen cell pellet of *E. coli* BL21 DE3 (pDB1013-5-4) (1 litre culture) prepared in a similar way to that described in Example 2b. was allowed to thaw at 4°C for 2 h and was then resuspended in 50 mM Tris/50 mM NaCl/1 mM EDTA/0.1 mM PMSF pH 8.0 (33 ml). The suspension was transferred to a 100 ml glass beaker and sonicated (Heat Systems - Ultrasonics W380; 70 Watts, 50 x 50% pulse, pulse time = 5 sec.). The sonicate was immediately centrifuged (6000g/4°C/10 min) and the supernatant was discarded. The pellet, containing the inclusion bodies, was resuspended in 20 mM Tris/8 M urea/50 mM 2-mercaptoethanol/1 mM EDTA/0.1 mM PMSF pH 8.5 (100 ml) and left static at room temperature (approx. 23°C) for 1 h. The resulting solution was centrifuged (approx. 2000 g at 4°C for 10 min) to remove any material that had failed to solubilise. The supernatant of this spin was retained at -40°C as the solubilised inclusion body product.

(d) **Purification of SCR 1 + 2 + 3 from the solubilised inclusion body**

A column (i.d., 16 mm; h, 10 mm) of S-Sepharose Fast Flow was prepared and connected into an FPLC (Pharmacia) system. The column was equilibrated with 20 mM Tris/8M urea/1 mM EDTA/50 mM 2-mercaptoethanol pH 8.5. 10 ml of

thawed, solubilised inclusion body, prepared as described in Example 2c, was applied to the column and washed through with equilibration buffer. The column was then developed with a linear gradient to 1M NaCl (in equilibration buffer) followed by rinses with 1M NaCl and 2M NaCl (also in equilibration buffer). All the chromatography was at 1.0 ml min⁻¹ and at room temperature.

Analysis by SDS PAGE/protein staining of the fractions collected during the chromatography indicated that virtually all the SCR 1 + 2 + 3 polypeptide had absorbed to the column and had been dissociated by the 1M NaCl - containing buffer. The appropriate fractions were stored at -40°C.

Subsequent assay for protein content of the peak fraction using the Bradford protein assay and a bovine serum albumin standard showed it contained 2.8 mg protein.

(e) Folding

S-Sepharose - purified SCR 1 + 2 + 3 that had been purified in a similar way to that described in Example 2d and stored at -40°C was thawed and 0.4 ml buffer-exchanged into 0.05 M formic acid using Sephaex G25 (P10). The absorbance at 280 nm of the buffer-exchanged solution was determined as 0.52, and, using $\epsilon = 34000$ and appropriate correction factors for dilution, the protein concentration of the original preparation (prior to buffer-exchange) was calculated to be 0.6 mg/ml.

Based on this figure, 1.7 ml S-Sepharose - purified protein was diluted with 0.85 ml 20 mM Tris/8M urea/50 mM 2-mercaptoethanol/1M NaCl pH 8.5 to give a 0.4 mg/ml solution, on which the folding was carried out.

Folding was effected by a series of dilutions, using cold diluent at all times.

At t=0 h, 0.8 ml SCR 1 + 2 + 3 (0.4 mg/ml) was added to 0.8 ml 20 mM Tris/1M urea/5 mM EDTA/3 mM 2-mercaptoethanol pH 8.0 ('diluent') in a 30 ml polystyrene universal container. The solution was mixed thoroughly by gentle swirling and left static, capped, in a cold room (approx. 2 to 3°C).

At 1 h, 1.6 ml diluent was added and mixed.

At 2 h, 3.2 ml diluent was added and mixed.

At 4 h, 6.4 ml diluent was added and mixed.

The solution was left a further 20 h in the cold room, then ultrafiltered (YM5, Amicon Ltd) to approx. 1.4 ml. This was buffer-exchanged into 0.1M NH₄HCO₃ (2.5 ml) using Sephadex G25 (PD 10) in the cold room. The eluate was aliquoted and was stored at -40°C or lyophilised.

The product containing SCR 1 + 2 + 3 was analysed by SDS PAGE, followed by protein staining. In both non-reduced and reduced (with 2-mercaptoethanol 5% (v/v)) samples there was a single major band. The molecular weight of the reduced band, compared to reduced protein standards of known M_r, was approx. 24,000. The

non-reduced protein (band) had a slightly faster mobility than the reduced protein (band).

The product was analysed in a functional haemolytic assay utilising antibody-sensitized sheep erythrocytes (Method (xvi)). The product showed concentration-dependent inhibition of the complement - mediated lysis of the erythrocytes with an
5 IH50 around 0.4 µg/ml.

(f) **Folding**

Preparation, folding, processing and analysis were carried out exactly as described in Example 2e except

- 10 (1) the diluent for the folding was 20 mM ethanolamine (pH 10.0)
(2) the folded solution was ultrafiltered to a final volume of 1.55 ml, and
(3) the IH50 figure was determined as about 0.6 µg/ml.
(4) the recovery of product was approx. 100 per cent.

(g) **Determination of N-terminal sequence of SCR 1 + 2 + 3**

- 15 1ml samples of growing *E.coli* BL21 (DE3) containing plasmid pDB1013-5-4 were removed 3 hours post-induction with 1 mM IPTG as described in Example 2b. These samples were spun in an eppendorf centrifuge and the resultant pellets each resuspended in 200µl of reducing buffer (100 mM Tris pH6.8/200 mM dithiothreitol/4% (w/v) SDS/2% (w/v) bromophenol blue and 20% (v/v) glycerol and
20 boiled for 5 minutes. 25µl samples were applied to a 14% polyacrylamide gel. When the electrophoresis was complete the proteins were transferred to a ProBlott membrane (Applied Biosystems) using a Sartoblot electroblotting apparatus (Sartorius) at 0.8 mA/cm² for 1 hour 40 mins using CAPS (3-[cyclohexylamino]-1-propanesulphonic acid) transfer buffer. After transfer the ProBlott membrane was
25 stained (0.1% (w/v) Coomassie Blue R-250/40% (v/v) methanol/1% v/v acetic acid) for 20 seconds and destained using 50% (v/v) methanol. A band corresponding to a M_r approx 23,000 protein was excised and the N-terminal sequence determined using a Blott cartridge in an Applied Biosystems 477A Protein Sequencer.

- 30 The sequence of the first 20 amino acids was found to agree with the predicted sequence except that residue 3 could not be identified by the sequencing protocol used.

Example 3 Expression and purification of SCR 1 + 2 + 3 from a fermentation vessel

- 35 (a) **Fermentation of *E. coli* harbouring the plasmid pDB1013-5-4**

E. coli BL21 (DE3) : pDB1013-5-4 was recovered from storage in liquid nitrogen by thawing a vial containing 1 ml of the culture and this was used to inoculate 100 ml of seed medium (NCYZM) containing ampicillin at 75 µg/ml. The

primary and secondary seed stage fermentations were carried out in plain 500 ml shake flasks batched with 100 ml aliquots of NCYZM medium. The primary and secondary seed fermentation conditions were as follows : 37°C, 230 rpm on an orbital shaker with a 50 mm throw. The primary seed incubation time was 2 hours.

- 5 The primary seed culture was used to inoculate secondary seed fermentation medium at 0.1% (v/v). The secondary seed was incubated for 14.5 hours.

- Two 15 litre Biolafitte fermenters were each batched with 10 litres of NCYZM medium and 0.01% (v/v) Dow Corning DC1510 antifoam. The vessels plus media were sterilised using steam to 121°C for 45 minutes. Ampicillin
10 sterilised by microfiltration (0.2 µm) was added aseptically to the vessel media to a final concentration of 150 µg/ml. The fermenters were inoculated at a level of 3% (v/v) from pooled secondary seed culture. The final stage incubation conditions were 37°C, agitator 400 rpm, airflow 5 l/min (0.5 vvm). The final stage fermentations were sampled aseptically pre-inoculation, at 0 hours and thence every half hour. The
15 samples were monitored for increases in optical density (600 nm). When the OD600 was ≥ 0.5 , IPTG was added to give a final concentration of 1 mM. The fermentations were incubated for a further 3 hours.

The cells were recovered by centrifugation using 7000 g for 25 minutes. The total cell yield (wet weight) was 49.8 grammes.

- 20 (b) **Inclusion body isolation.**

Inclusion bodies from 23 g (wet weight) cell pellet were isolated and solubilised essentially as described in Example 2.

- (c) **Purification of denatured SCR 1 + 2 + 3**

- The volume of solubilised inclusion body from Example 3b was approx. 800
25 ml. To this viscous solution was added SP-Sepharose FF (100 ml gel bed, water washed and suction dried). The mixture was swirled vigorously and left static for 1h at room temperature. The supernatant was decanted and stored at -40°C. The remaining slurry was resuspended to a uniform suspension and poured into a glass jacket (i.d., 41.5 mm) and allowed to settle into a packed bed. This packed bed was
30 connected into a low pressure chromatography system at 4°C and equilibrated with 0.02M Tris/8M urea/0.05M 2-mercaptoethanol/0.001 M EDTA pH 8.5. When the A₂₈₀ of the eluate had minimised, the buffer was changed (step-wise) to the equilibration buffer additionally containing 1 M NaCl. A single A₂₈₀ peak was eluted, in a volume of 90 ml (equivalent to approx. 1 V_t). The solution was clear and
35 colourless and was estimated, by A₂₈₀ determination of a buffer-exchanged sample (using an $\epsilon = 25,000$), to contain about 300 mg target protein. By SDS PAGE followed by protein stain the target protein was the major band present. The 90 ml product was stored at -40°C.

(d) **Folding and further purification.**

18 ml of the above product (nominal 60 mg) was diluted with 12 ml 0.02 M Tris/8M urea/1 M NaCl/0.05M 2-mercaptoethanol pH 8.5. The product (30 ml) was added as 5 ml aliquots at 1 min intervals to 930 ml freshly prepared, cold 0.02 M ethanolamine/0.001 M EDTA, with swirling, and left static for 1 h/4°C. Then reduced glutathione was added to 1 mM (by addition of 9.6 ml 0.1M GSH) and oxidised glutathione was added to 0.5 mM (by addition of 9.6 ml 0.05M GSSG). The solution was clear and was left static in the cold for approx. 70 h. The solution was then ultrafiltered using a YM10 membrane to a final retentate volume of about 10 ml; this retentate was cloudy. It was mixed with 90 ml 0.1 M NaH₂PO₄/1 M (NH₄)₂SO₄ pH 7.0 (Buffer A) at room temperature and then centrifuged at 4000 rpm for 20 min. The supernatant was decanted and SCR 1 + 2 + 3 protein isolated by chromatography of the supernatant on Butyl Toyopearl 650 S.

The column of Butyl Toyopearl (V_t ~ 12 ml) was equilibrated with Buffer A. The 100 ml supernatant was applied to the column and the column washed with Buffer A. It was then developed with a linear gradient of 100% Buffer A to 100% 0.1 M NaH₂PO₄ pH 7.0. All the chromatography was at room temperature at approx. 30 cmh⁻¹.

A major A₂₈₀ peak was eluted during the gradient. Fractions spanning the peak were analysed by SDS PAGE followed by protein stain. The most concentrated fractions of the peak contained essentially pure SCR 1 + 2 + 3 and were active in the haemolytic assay (Method (xvi)) (IH₅₀ ~ 0.3 µg/ml). They were stored at -40°C.

Example 4 Formulation of Butyl Toyopearl purified SCR 1 + 2 + 3.

Batches of SCR 1 + 2 + 3 that had been expressed, folded and purified in similar ways to batches described in Examples 2 and 3 and further purified by ammonium sulphate treatment and Butyl Toyopearl chromatography essentially as described in Example 3d were formulated into a useable product as follows.

Three such Butyl Toyopearl products were pooled to give a volume of about 31 ml. All 31 ml were buffer-exchanged into 0.05 M formic acid (prepared using 0.2 µm-filtered 'MilliQ' water) using a column of Sephadex G25. All the chromatography was at 50 cmh⁻¹ at 4°C. The eluate from the column was monitored at 280 nm and the V₀ fraction was collected as a single fraction. The bulk of this fraction was lyophilised in aliquots.

Analysis of the V₀ pool prior to lyophilisation by both SDS PAGE/stain and C8 reverse phase HPLC showed it to be essentially pure target protein. The pool demonstrated anti-haemolytic activity (IH₅₀ approx. 0.3 µg/ml) and the endotoxin content was low (< 1 ng/mg).

One of the lyophilised aliquots was shown to be soluble at 10 mg ml^{-1} in phosphate-buffered saline and showed complement inhibitory activity in the haemolytic assay (Method xvi); the IH_{50} was $0.3 \mu\text{g/ml}$.

Another of the lyophilised aliquots was examined to determine the disulphide bridge pattern. All six correct (as predicted on the basis of a consensus SCR motif) disulphides were detected.

appears

Example 5 Effect of SCR 1+2+3 on IgG-mediated activation of the classical pathway of complement, as measured by C3a release

Inhibition of heat aggregated IgG activated serum was carried out as described in Method (xvii). Heat aggregated IgG activates the classical pathway of complement. Different concentrations (typically $4 - 1000 \mu\text{g/ml}$) of inhibitor were incubated with serum in the presence of heat aggregated IgG and the % inhibition of activation at each concentration was determined. The IC_{50} of SCR 1+2+3 was determined as approximately $22 \mu\text{g/ml}$ indicating that SCR 1+2+3 can inhibit the classical pathway of complement.

Example 6 Effect of SCR 1+2+3 on zymosan A-mediated activation of the alternative pathway of complement, as measured by following C3a release.

Inhibition of zymosan A activated serum was carried out as described in Method (xvii). Different concentrations of SCR 1+2+3 (typically in the range $1 - 1000 \mu\text{g/ml}$) were incubated with serum in the presence of zymosan A and the % inhibition of activation at each concentration was determined. From several different experiments the IC_{50} was determined as $20 - 40 \mu\text{g/ml}$ indicating that SCR 1+2+3 can inhibit the alternative pathway of complement.

Example 7 Effect of SCR 1+2+3 on zymosan A-mediated activation of the alternative pathway of complement, as measured by C5a release.

Inhibition of zymosan A activated serum was carried out as described in Method (xvii) and assayed as described in Method (xviii). Different concentrations of SCR 1+2+3 (typically in the range $4 - 700 \mu\text{g/ml}$) were incubated with serum in the presence of zymosan A and the % inhibition of activation at each concentration was determined. From several different experiments the IC_{50} was determined as approximately $20 - 30 \mu\text{g/ml}$, indicating that SCR 1 + 2 + 3 can inhibit the alternative pathway of complement.

Example 8 Endotoxin content determination of purified, folded and formulated SCR 1 + 2 + 3

A batch of final product SCR 1 + 2 + 3 was prepared essentially as described in Example 4 above and was measured for endotoxin content using a method based on the gel-clot reaction of limulus amoebocyte lysate (LAL) (Atlas Bioscan Ltd.). The sensitivity of the assay was 0.125 EU/ml and this was checked by titration against a doubling dilution series prepared from standard *E. coli* endotoxin supplied with the LAL kit.

10-fold dilutions of ~ 1.3 mg/ml SCR 1 + 2 + 3 protein stock were tested in quadruplicate for their effect on LAL by adding 10µl of sample to 10µl LAL. After 1h at 37°C the mixtures were tested for either clotting or remaining liquid. (Solutions that contain at least 0.125 EU endotoxin will clot this LAL preparation.) After taking into account the results of simultaneous tests designed to test for interference, it was concluded that the endotoxin content of the SCR 1 + 2 + 3 protein preparation was < 12.5 EU/ml, equivalent to approx. < 1 ng/mg protein.

Example 9 Expression, folding, purification, and formulation of MR122 -> K196 of CR-1 (SCR 3)

General Points

The sequence for SCR 3 corresponding to amino acid 122 and ending at amino acid 196 of mature human complement receptor 1 was designed such that the 5' end of the gene contained an *NdeI* restriction endonuclease site. This site comprises an ATG start codon to give the initiating methionine required for the start of mRNA translation and allows the placement of the gene an optimum distance from the Shine-Dalgarno ribosome binding site of pBROC413. This codon is followed immediately by the gene coding for SCR 3 starting with arginine 122 of mature human complement receptor 1. The 3' end of the gene finishes with a codon for lysine 196 followed by two stop codons followed by a *HindIII* site.

The DNA coding for SCR3 was modified for optimum codon usage in *E. coli* as described in the methods. The gene was also altered to incorporate unique restriction endonuclease sites. This was carried out in the following way. Restriction endonucleases that do not cut pBROC413 and were commercially available were identified. The DNA sequence of these restriction endonuclease sites was then translated into all three reading frames and the codon usage examined. Sites that contained codons that are rarely used by *E. coli* were discarded. The remaining sites were examined for their translated sequence and if that sequence matched with SCR 3, the restriction site was incorporated into the sequence.

(a) Construction of plasmid pBROC435 encoding SCR 3

The construction of pBROC435 is described in Example 2a

(b) **Expression of SCR 3 from pBROC435**

pBROC435 was transformed by electroporation into *E.coli* BL21(DE3) and resulting colonies analysed by restriction digestion of mini-plasmid DNA

- 5 preparations. Single colonies were inoculated into universals containing 10 ml of L broth or NCYZM medium and 50 - 75 µg/ml ampicillin and allowed to grow overnight at 37°C, 220 r.p.m. Typically 4ml of overnight cultures were used to inoculate each of 2 L conical flasks containing 500 ml of NCYZM medium, 150 µg/ml ampicillin; cultures were grown at 37°C, 230r.p.m. until A₆₀₀ was 0.5
- 10 absorbance units. Cultures were induced with 1 mM IPTG and allowed to grow a further 3 hours under the same conditions. The cultures were centrifuged (approx, 8000g/10 min) and the supernatants discarded. The cell pellets were stored at -40°C.

(c) **Isolation of solubilised inclusion bodies**

- The frozen cell pellet of *E. coli* (from 3 l growth culture in NCYZM) described in Example 9b was allowed to thaw at room temperature for 2 h and was
- 15 then resuspended in 50 mM Tris/50mM NaCl/1mM EDTA/0.1 mM PMSF pH 8.0 (90ml). The suspension was transferred to a 200 ml glass beaker and sonicated (Heat Systems - Ultrasonics W380; 70 Watts, 50 x 50% pulse, pulse time = 5 sec.). The sonicate was immediately centrifuged (6000g/4°C/10 min) and the supernatant was
- 20 discarded. The pellet, containing the inclusion bodies, was resuspended in 20 mM Tris/8 M urea/50 mM 2-mercaptoethanol/1 mM EDTA/0.1mM PMSF pH 8.5 (300ml) with gentle pipetting to mix. After mixing, the solution was left static at room temperature (approx. 23° C) for 1 h. The resulting solution was centrifuged
- 25 (2000 g at 4° C for 10 min) to remove material that failed to solubilise. The supernatant of this spin was retained at -40° C as the solubilised inclusion body product.

(d) **Purification of SCR3 from the solubilised inclusion body**

- A column (i.d., 32 mm; h, 32mm) of Q-Sepharose Fast Flow (Pharmacia) was prepared and equilibrated with 20 mM Tris/8M urea/50 mM 2-mercaptoethanol pH
- 30 9.0. 200 ml of thawed, solubilised inclusion body, prepared as in Example 9c, was applied to the column and washed through with equilibration buffer. The column was connected to an FPLC system and developed via a stepwise gradient of 0, 1.0, 2.0M NaCl (also in equilibration buffer). All chromatography was at 2.0 ml min⁻¹ and at room temperature.

- 35 Analysis by SDS PAGE/protein staining of the fractions collected during the chromatography indicated that virtually all the SCR3 did not bind to the column. Many other proteins had absorbed to the column however and had been dissociated by the 0.1M and 1M NaCl - containing buffers. The purity of SCR3 in the

unadsorbed fraction was estimated to be about 80%.

(e) **Folding of SCR3**

Q-Sepharose-purified SCR3 that had been purified as described in Example 9d and stored at -40°C was thawed and was folded by a series of dilutions, using cold diluent. At $t=0$, 100 ml 20 mM Tris/1M urea/5mM EDTA/3mM 2-mercaptoethanol pH 8.0 (diluent) were added to 100ml SCR3. At this stage the solution was turbid in appearance. The solution was mixed thoroughly by gentle swirling and left static, capped, in a cold room ($2-3^{\circ}\text{C}$). At 1 h, 200ml diluent was added and mixed, final volume = 400 ml. At 2 h, 400ml diluent was added and mixed, final volume = 800ml. At 4 h 800ml of diluent was added and mixed, final volume = 1.6L. The solution was left for a further 20 h in the cold room. The solution now appeared clear, and it was stored at -40°C in aliquots.

(f) **Formulation of SCR3**

50 ml of SCR3 prepared as in Example 9e were thawed and ultrafiltered to 3.5 ml using a 2000 Da cut-off membrane (Amicon). 2.5 ml of the concentrate was buffer-exchanged into 0.1M NH_4HCO_3 (3.0 ml) using Sephadex G25 (PD 10). Subsequent analysis for protein content using the molar extinction coefficient of 11000 showed this sample contained approx 0.24mg/ml.

Analysis of this material by SDS PAGE/protein staining indicated that the protein was about 80% pure. Samples reduced with 2-mercaptoethanol had a lower electrophoretic mobility suggesting the presence of disulphide bonds in SCR3.

Analysis of this sample in the haemolytic assay (Method (xvi)) showed it had an IH_{50} of approx. 10 - 20 $\mu\text{g/ml}$.

(g) **Determination of N-terminal sequence of expressed SCR3**

200 μl SCR3 prepared and formulated in 0.1M NH_4HCO_3 as in Example 9f was precipitated with 800 μl cold acetone in a cardice/ethanol bath for 60 mins. The sample was then spun in an Eppendorf centrifuge (approx 10,000g/20 mins) and the resultant pellet resuspended with heating in sample buffer containing 5% (v/v) 2-mercaptoethanol. 30 μl samples were electrophoresed on a 4 to 20% SDS-containing polyacrylamide gradient gel. When the electrophoresis was complete the proteins were transferred to a ProBlott membrane (Applied Biosystems) using an electroblotting apparatus at 200mA for 2 h using CAPS in 10% methanol/90% H_2O (v/v) transfer buffer. After transfer the ProBlott membrane was stained (0.1% (w/v) Coomassie Blue), destained, rinsed and air dried according to the manufacturer's instructions. Sections of the membrane were excised and used for N-terminal sequencing.

The sequence of the first 20 amino acids of the major band was as expected for SCR3 with the exception of residue 5, which could not be identified.

(h) **Preparation, folding and formulation of SCR3**

Preparation and folding were carried out exactly as described in Example 9a-9e. 400 ml of folded SCR3 was ultrafiltered through a 30 KDa cut-off filter (Amicon) at 4^o C. Samples of the ultrafiltrate were processed in two ways.

- 5 1. 50 ml were ultrafiltered using a 2 KDa cut-off membrane to a final volume of 3.5ml and buffer-exchanged into 0.05 M formic acid (6.7 ml) using Sephadex G25 (PD10) columns. The total amount of SCR3 estimated by the absorbance at 280 nm was 0.6mg. Analysis by SDS PAGE/protein staining indicated that the protein had a purity of about 95%. The sample was freeze-dried
10 and stored at -40^oC.
2. 100ml of the ultrafiltrate were adjusted to pH 5.5 with HCl. The sample was applied to a Mono S column (1ml) at 1.5 ml min⁻¹ and washed through with equilibration buffer (20mM Tris.HCl pH 5.5). The column was then developed with a step gradient of 0.1, 1.0 and 2.0M NaCl (also in equilibration buffer). All
15 remaining chromatography was at 1.0 ml min⁻¹ and at room temperature.

Analysis by SDS PAGE/protein staining of the fractions collected during the chromatography demonstrated that the major band dissociated at 1M NaCl contained SCR3 at about 95% purity.

20 **Example 10 Expression, folding, purification and formulation of MR122-S253 of CR-1 (SCR 3 + 4)**

(a) **Construction of plasmid pDB1019 encoding SCR 3 + 4**

The DNA coding for SCR 3 + 4 was constructed from the plasmids pBROC435 (Example 2) and pDB1018-1 (Example 11) which carry the genes coding
25 for SCR 3 and SCR 1 + 2 + 3 +4 respectively. The SCR 4 coding unit was excised from pDB1018-1 and ligated onto the end of the SCR 3 coding unit in pBROC435.

pDB1018-1 was digested with *SpeI* and *HindIII* and separated on a 1 % agarose gel. The band which codes for SCR 4 (~ 245 bp) was excised from the gel and purified using the QIAEX extraction kit. Plasmid pBROC435 was also cut with
30 *SpeI* and *HindIII*, separated on 1 % agarose, excised from the agarose and purified with the QIAEX kit. The SCR 4 coding DNA was then ligated into the cut pBROC435 plasmid to give pDB1019. This DNA was used to transform *E. coli* HB101 made competent with CaCl₂. Transformants were analysed by restriction analysis using *EcoRI* and *HindIII*. Clones carrying the correct sized insert were used
35 for expression studies.

(b) **Expression of SCR 3 + 4 from pDB1019-1C**

pDB1019 was transformed into *E. coli* BL21(DE3) made competent with CaCl₂ and the resulting colonies were analysed by restriction digestion of mini-

plasmid DNA preparations. Plasmid pDB1019-1C was identified as carrying the correct sized insert. Single colonies of *E. coli* BL21(DE3) carrying pDB1019-1C were inoculated into ten universals containing 10 mls of NCYZM medium and 75 µg/ml ampicillin and allowed to grow overnight at 37°C, 240 r.p.m. The overnight
5 cultures were then used to inoculate eight 2 L conical flasks (5 ml/flask) containing 500 ml of NCYZM medium, 150 µg/ml ampicillin. Cultures were grown at 37°C, 240 r.p.m. until A₆₀₀ was 0.5 absorbance units. At this point cultures were induced with 1 mM IPTG and allowed to grow a further 3 hours under the same conditions. The cultures were centrifuged (approx. 8000g/ 10 mins) and the supernatants were
10 discarded. The cell pellets were stored at -40°C.

(c) **Isolation, purification, folding and formulation of SCR 3 + 4**

The methods used generally follow those described earlier for the preparation of SCR 1 + 2 + 3.

Isolation of solubilised inclusion bodies from cell pellet derived from 21
15 culture was carried out as described in Example 2c. The volume of solubilisate was 200 ml.

Some of the contaminating (host) *E. coli* proteins were removed from the preparation by adsorption onto S-Sepharose, either in a batch process or by column chromatography, using systems similar to those described in Example 2d. The
20 protein present in the unadsorbed fractions was shown by SDS PAGE/stain to contain significant amounts of SCR 3 + 4 protein. About half of these fractions were ultrafiltered using a YM1 (Amicon) membrane to approx. 35 to 40 ml. This retentate was estimated to contain about 0.3 mg protein/ml (based on A₂₈₀ determination of a buffer-exchanged sample, using $\epsilon = 21,000$). 10.5 ml of the retentate was mixed with
25 325 ml cold 20 mM ethanolamine and left static at 4°C for 1 hour. Then reduced glutathione was added to 1 mM (by addition of 3.4 ml 100mM GSH) and oxidised glutathione was added to 0.5mM (by addition of 3.4ml 50mM GSSG). The solution was mixed and left static at 4°C for ~ 72 h. The solution was clear. The solution was then ultrafiltered using a YM1 membrane to a retentate of 5 ml. The retentate
30 was divided in two and buffer-exchanged into either 20 mM ethanolamine or 50 mM formic acid using Sephadex G25 (PD10 columns).

Analysis of the formic acid SCR 3 + 4 product by reverse phase HPLC and by SDS PAGE followed by protein staining showed only one major protein species (> 90% pure). The protein concentration was estimated to be 0.3 mg/ml using A₂₈₀
35 determinations. The product was active in the haemolytic assay (Method (xvi)); the IH50 value was approx. 30 µg/ml

Example 11 Construction, expression, folding, purification and formulation of MQ1-S253 of CR-1 (SCR 1 + 2 + 3 + 4)

General points

Two constructs were prepared by making a plasmid encoding SCR 1+2, incorporating SCR3 and finally adding SCR4. The two constructs encoded consensus SCR1 to 4 and the R235H mutation of SCR1 to 4 (Example 12).

A plasmid containing the SCR 1 + 2 + 3 + 4 coding unit was constructed by adding the DNA encoding SCR 4 onto the construct coding for SCR 1 + 2 + 3 (Example 2). For convenience of DNA manipulation, the SCR 4 DNA coding unit was made by synthesising the DNA encoding the last 17 amino acids of SCR 3 followed by the DNA coding for the linker region followed by SCR 4. This DNA started at the *SpeI* site of the SCR 1 + 2 + 3 coding construct which corresponds to T175 of mature CR-1 followed by the DNA coding for the linker region followed by SCR 4 ending on the codon corresponding to S253 followed by two stop codons and a *HindIII* site. As for the previous constructs the DNA encoding SCR 4 was altered for optimised codon usage and restriction sites as previously described in Example 1. This unit of DNA was ligated to the plasmid coding for SCR 1 + 2 + 3 which had been cut with *SpeI* and *HindIII* to give a construct coding for SCR 1 + 2 + 3 + 4.

(a) Construction of plasmid pDB1018 encoding SCR 1 + 2 + 3 + 4

Oligonucleotides (Table 1; oligos 21 - 26 coding for SCR4) were synthesised as 3 complementary pairs of 68-90 mers that could be ligated in a unique fashion via complementary 8 bp overhangs between the pairs of oligonucleotides. The 3 complementary pairs of oligonucleotides were designated Pair E (oligos 21, 22), Pair F (oligos 23, 24) and Pair G (oligos 25, 26). Pair E which corresponds to the 5' end of the gene contained a *SpeI* restriction site overhang and Pair G contained a *Hind III* restriction site overhang at the 3' end. All oligonucleotides apart from 22 and 24 were purified by electrophoresis through a denaturing polyacrylamide gel followed by reverse phase chromatography (C₁₈). Oligonucleotides 22, 23, 24 and 25 were kinased before annealing to their complementary oligonucleotides. The oligonucleotides were ligated pair E to pair F to pair G to form the gene coding for part of SCR3 and the whole of SCR4 which for convenience will be called the SCR4 gene in the subsequent text.

The DNA coding for SCR4 was initially amplified by PCR using two oligonucleotides (Table 1; oligos 17 and 18) complementary to the two strands of DNA. Both oligonucleotides contained 5' unmatched ends that contained 6bp of random sequence followed by the sequence of either *SpeI* (oligo 17) or *Hind III* (oligo 18) restriction sites followed by 18 bp complementary to the gene. Following PCR a band of approximately 250 bp was visualised on horizontal agarose gel

electrophoresis, which was excised and purified on DEAE NA45 membranes. This DNA was used for a second PCR amplification using nested primers that had been moved inwards by four nucleotides at their 5' ends (Table 1; oligo 19, oligo 20). These oligo's incorporated the *Spe*I and *Hind*III restriction sites but now only had 2 nucleotides beyond the end of each restriction site. Following PCR a band of approx. 250 bp was visualised on horizontal agarose gel electrophoresis. This band was excised and purified using the QIAEX agarose gel extraction kit.

The DNA for SCR 4 was blunt-end ligated to itself following kinasing. The multimers formed were visualized by horizontal agarose gel electrophoresis and the bands excised and purified using the QIAEX agarose gel extraction kit. The DNA was then cut with *Spe* I and *Hind* III and ligated into pDB1013-5-4 that had been cut with the same enzymes to produce pDB1018 (Fig.3). The vector was transformed into *E.coli* HB101 made competent with calcium chloride. Mini-plasmid preparations were made and plasmid DNA analysed by digestion with *Nde* I, *Hind* III, *Stu* I, *Spe* I and *Kpn* I. The plasmids with the correct restriction maps were analysed by DNA sequencing of both strands across the gene encoding SCR4. Two plasmids were selected for further study. pDB1018-1, which encoded MQ1-S253 (consensus SCR1 to 4) and pDB1018-6, which encoded the R235H mutant of MQ1-S253. The amino acid sequences of the two polypeptides encoded by pDB1018-1 and pDB1018-6 are shown in Table 2.

Taking the first residue as being the A of the ATG initiating codon, DNA sequencing revealed that residue 600 of pDB1018-6 had been altered from G -> A. This is a silent mutation and does not alter the amino acid at this position.

(b) **Expression of MQ1-S253 from pDB1018-1**

pDB1018-1, constructed as described in Example 11a, was transformed into calcium chloride competent *E.coli* BL21(DE3). Single colonies were inoculated into universals containing 10ml of NZCYM medium and 75 µg/ml ampicillin and allowed to grow overnight at 37°C, 230 r.p.m. 3ml of overnight culture were used to inoculate each of 8 x 2 litre conical flasks containing 500ml of NZCYM medium, 150 µg/ml ampicillin; cultures were grown at 37°C, 230 r.p.m. until A₆₀₀ reached 0.5 absorbance units. The cultures were induced with 1mM IPTG and allowed to grow for a further 3 hours under the same conditions. The cultures were centrifuged (approx. 7000g/10 mins/4°C) and the supernatants discarded. The cell pellets were stored at -40°C.

(c) **Isolation of solubilised inclusion bodies**

The frozen cell pellets of *E.coli* BL21(DE3) (pDB1018-1) each equivalent to 1 litre of culture prepared as described in Example 11b were allowed to thaw at 0-4°C over 2 hours. The pellets were resuspended in 50mM Tris/50mM NaCl/1mM

EDTA/0.1mM PMSF pH 8.0: 30ml for each litre pellet. Each suspension was transferred to a 100ml glass beaker and sonicated (Heat systems - Ultrasonics W380; 70 Watts, 50 x 50% pulse, pulse time = 5 seconds). The sonicates were pooled and immediately centrifuged (6,000 g/ 4°C/10 mins) and the supernatant discarded. The pellet containing the inclusion bodies was resuspended in 20 mM Tris/8 M urea/50mM 2-mercaptoethanol/1mM EDTA/0.1mM PMSF pH 8.5 (400 ml), thoroughly mixed and left static at room temperature (approx. 23°C) for 1 hour.

(d) **Purification of MQ1-S253 from the solubilised inclusion body.**

30ml of S-Sepharose FF that had been washed with deionised water and suction dried was added to the inclusion body solution described in Example 11c, and vigorously shaken for 30 seconds. The S-Sepharose mixture was left static at room temperature (23°C) for 1.5 hours and then the supernatant was discarded. The remaining slurry was packed into a column (id, 4.1cm). The column was equilibrated using 20mM Tris/8M urea/50mM 2-mercaptoethanol/1mM EDTA/0.1mM PMSF pH 8.5 at 60 cmh⁻¹, 4°C. MQ1-S253 protein was eluted using the equilibration buffer containing 1M NaCl. Analysis by SDS PAGE/protein staining of the fractions collected during the chromatography indicated that virtually all the target protein had adsorbed to the column and had been dissociated by the 1M NaCl wash. The appropriate fraction was stored at -40°C.

(e) **Folding and formulation**

Based on a molar extinction coefficient of 25,000 and A₂₈₀ values determined in 50mM formic acid, 60 mg of the S-Sepharose purified unfolded protein described in Example 11d was folded and formulated as follows :-

8.0ml of solution (equivalent to 60 mg protein) was diluted with 22ml cold 20mM Tris/8M urea/50mM 2-mercaptoethanol/1M NaCl/1mM EDTA/0.1mM PMSF pH8.5, to give 30ml of a 2.0 mg/ml solution. The 30ml was diluted rapidly with constant stirring into 930ml cold (0-4°C) freshly prepared 20mM ethanolamine. The solution was left static at 0-4°C for 1 hour. Reduced glutathione was added to 1mM (by addition of 9.6ml of 100mM stock) and then oxidised glutathione was added to 0.5mM (by addition of 9.6ml of 50mM stock). The solution was left static at 0-4°C for a further 48 hours and then ultrafiltered using a stirred cell (Amicon) and a YM10 membrane (Amicon, nominal 10,000 Da molecular weight cut-off) to approx. 29ml. The ultrafiltered retentate was buffer exchanged into 50 mM formic acid using Sephadex 25 (i.d., 26mm; h, 245mm Vt, 123ml) and a flow rate of 50 cmh⁻¹ to a final volume of 40ml. Using a molar extinction coefficient of 25,000 for the protein 51mg of protein was recovered. The purified protein gave an IH₅₀ value (see Method xvi) of approximately 2 µg/ml.

(f) **Further purification and formulation of SCR1 + 2 + 3 + 4.**

Folded SCR1 + 2 + 3 + 4 (nominal 25mg) in 50mM formic acid prepared essentially as described in Example 11e was lyophilised. The lyophilisate was resolubilised in 20mM ethanolamine (10ml) to give a cloudy solution. The 10ml
5 were then added to 90ml 0.1M NaH₂PO₄/1M (NH₄)₂SO₄ pH 7.0, thoroughly mixed, and then clarified by centrifugation (4000 rpm/20 min). The supernatant (100ml) was decanted and was chromatographed on Butyl Toyopearl (exactly as described for SCR1 + 2 + 3 in Example 3d). The peak A₂₈₀ fractions, eluting at about 100% of the 1M NaCl-containing buffer, were pooled and buffer-exchanged using Sephadex
10 G25 into 50mM formic acid. The V₀ pool (29.5ml) was lyophilised in aliquots.

The purity of the protein was assessed by SDS PAGE followed by protein staining and by C8 reverse-phase HPLC; the protein was estimated to be >95% pure. One of the lyophilised aliquots was resolubilised to 4mg protein/ml in 0.1M Hepes/0.15M NaCl pH7.4. The product showed activity in the haemolytic assay
15 (Method (xvi)); the IH50 was calculated to be 0.3 µg/ml.

Another of the lyophilised aliquots was examined to determine the disulphide bridge pattern using proteolytic digestion and peptide identification by amino acid sequencing. All eight correct (as predicted on the basis of a consensus SCR motif) disulphides were detected.

20

Example 12 Expression, isolation, folding and formulation of purified MQ1-S253 (R235H)

(a) **Expression of MQ1-S253 (R235H)**

pDB1018-6 (prepared as described in Example 11a) was transformed into
25 calcium chloride competent *E. coli* BL21(DE3). Single colonies were inoculated into universals containing 10 mls of NCYZM medium and 50 µg/ml ampicillin and allowed to grow overnight at 37°C, 220 r.p.m. The overnight cultures (approx. 3ml) were used to inoculate each of 2 l conical flasks containing 500 ml of NCYZM medium, 150 µg/ml ampicillin; cultures were grown at 37°C, 220 r.p.m. until A₆₀₀
30 was 0.5 absorbance units. Cultures were induced with 1 mM IPTG and allowed to grow a further 3 hours under the same conditions. The cultures were centrifuged (approx. 8000g/10 min/4°C) and the supernatants discarded. The cell pellets were stored at -40°C.

(b) **Isolation of solubilised inclusion bodies and purification of unfolded MQ1-S253 (R235H)**

35

Frozen cell pellet of *E. coli* BL21 DE3 (pDB1018-6) (2 litre culture) described in Example 12a was allowed to thaw at 4°C for 2 h and was then resuspended in 50 mM Tris/50 mM NaCl/1 mM EDTA/0.1 mM PMSF pH 8.0 (66

ml). The suspension was transferred to a 250 ml glass beaker and sonicated (Heat Systems - Ultrasonics W380; 70 Watts, 30 x 50% pulse time = 5 seconds). The sonicate was immediately centrifuged (6000g/4°C/10 min) and the supernatant was discarded. The pellet, containing the inclusion bodies, was resuspended by vigorous swirling in 20 mM Tris/8 M urea/50 mM 2-mercaptoethanol/1 mM EDTA/0.1 mM PMSF pH 8.5 (200 ml) and left static at room temperature (approx. 23°C) for 1.5h. Water-washed, suction-dried S-Sepharose (equivalent to approx. 25 ml packed bed volume) was added to the 200 ml solubilised inclusion body and the mixture swirled vigorously to wet the Sepharose beads thoroughly. The mixture was left static at room temperature for 1h. The supernatant (approximately 150 ml) was decanted and discarded. The slurry remaining was resuspended to a uniform suspension by swirling and then poured into a 32 mm (i.d.) glass jacket and allowed to settle. The gel bed was connected into a low pressure chromatography system and was equilibrated with 20 mM Tris/8 M urea/1mM EDTA/50mM 2-mercaptoethanol pH 8.5 at 4°C until the A₂₈₀ baseline stabilised. The column was then developed with equilibration buffer containing 1M NaCl. All the chromatography was at approx. 1 ml min⁻¹. Analysis by SDS PAGE/protein staining of the fractions collected during the chromatography indicated that most of the MQ1-S253 (R235H) polypeptide had adsorbed to the column and had been dissociated by the 1M NaCl - containing buffer wash and that the purity of the material was about 90%.

A sample of the pool was buffer-exchanged into 50mM formic acid using Sephadex G25 column to allow some assays to be carried out.

Amino acid analysis of the pool of the MQ1-S253 (R235H) - containing fractions gave a total protein content of about 120 mg.

25 (c) **Folding and formulation of SCR 1 + 2 + 3 (R235H)**

Based on A₂₈₀ values and a molar extinction coefficient of 25,000 for the protein in 50mM formic acid, 20 mg of the S-Sepharose-purified unfolded protein described in Example 12b was folded and formulated as follows.

30 5.2 ml protein solution (equivalent to 20 mg) was diluted with 4.8 ml cold 20 mM Tris/8M urea/50 mM 2-mercaptoethanol/1M NaCl pH8.5, to yield 10 ml of a 2.0 mg/ml solution.

35 The 10 ml was diluted rapidly with constant stirring into 310 ml freshly prepared, cold (approx. 0-4°C) 20 mM ethanolamine. The solution was left static at 0-4°C for 1 h. Then reduced glutathione was added to 1 mM (by addition of 2.5 ml 125 mM GSH). Then oxidised glutathione was added to 0.5 mM (by addition of 5.2 ml 50 mM GSSG). The solution was left static, in the cold room (~ 2-3°C), for a further 48 h. The solution was then ultrafiltered using a stirred cell and a YM10 (nominal 10,000 molecular weight out-off) membrane to approximately 2 ml. The

solution was clear. The ultrafiltration cell was washed with approximately 2 ml 20 mM ethanolamine and the wash and the ultrafiltered retentate were pooled to give a final volume of 3.7 ml.

2.2 ml of this solution was buffer-exchanged into 3.2 ml 50 mM formic acid using Sephadex G25 (PD10). The buffer-exchanged material was regarded as the product, and it was stored at -40°C. Analysis of an aliquot of the product showed it contained 1.6 mg protein/ml, that by SDS PAGE under non-reducing conditions a single major band of $M_r \sim 28,000$ was present and that N-terminal sequencing of the band (MQXNAPE) was consistent with the expected sequence. In addition the preparation gave an IH_{50} value (see Method (xvi)) of approximately 1 µg/ml.

IN THE FIGURES

Fig. 1 Plasmid pBROC413. *bla* indicates the ampicillin resistance gene, $\phi 10$ the T7 RNA polymerase promoter and rbs the ribosome binding site. Arrows for $\phi 10$ and *bla* give the direction of transcription. The polylinker site has been indicated. The plasmid is not drawn to scale and the size is approximate.

Figure 2 illustrates the construction from pDB1010-D11 and pBROC435 of plasmid pDB1013-5-4 coding for SCR 1 + 2 + 3. Plasmid sizes are approximate and are not drawn to scale.

Figure 3 illustrates the construction from pDB1013-5-4 of pDB1018 coding for SCR 1+2+3+4. Plasmid sizes are approximate and are not drawn to scale.

REFERENCES USED IN EXAMPLES OR GENERAL METHODS

1. Chen G-F. and Inouye M. (1990). Suppression of the negative effect of minor arginine codons on gene expression; preferential usage of minor codons within the first 25 codons of the *E.coli* genes. *Nuc Acids.Res.* 18 (6): 1465-1473.
2. Dacie, J.V. & Lewis S.M., (1975) *Practical haematology* Fifth Edition, Ed. Churchill Livingstone, Edinburgh & New York pp 3 - 4
3. Devereux, Haeberli and Smithies (1984). A comprehensive set of sequence analysis programmes for the vax. *Nuc Acids.Res.* 12(1): 387 - 395
4. Hourcade D., Miesner D. R., Atkinson J. P. & Holers V. M. (1988). Identification of an alternative polyadenylation site in the human C3b/C4b receptor (complement receptor type 1) transcriptional unit and prediction of a secreted form of complement receptor type 1. *J. Exp. Med.* 168 1255 - 1270
5. Low B. (1968). *Proc.Natl.Acad.Sci.USA* 60: 160
6. Sambrook J., Fritsch E.F. and Maniatis J. (1989). *Molecular Cloning: A Laboratory Manual* 2nd Edition. Cold Spring Harbour Laboratory Press

7. Studier F.W. and Moffat B.A. (1986). *J.Mol.Biol* 189: 113
8. Tabor S. (1990). Expression using the T7 RNA polymerase/promoter system. In Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, .D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl, eds) pp. 16.2.1 - 16.2.11 Greene Publishing and Wiley-Interscience, New York

TABLE 1

	OLIGO 1 = SEQ ID NO: 1
10	OLIGO 2 = SEQ ID NO: 2
	OLIGO 3 = SEQ ID NO: 3
	OLIGO 4 = SEQ ID NO: 4
	OLIGO 5 = SEQ ID NO: 5
	OLIGO 6 = SEQ ID NO: 6
15	OLIGO 7 = SEQ ID NO: 7
	OLIGO 8 = SEQ ID NO: 8
	OLIGO 9 = SEQ ID NO: 9
	OLIGO 10 = SEQ ID NO: 10
	OLIGO 11 = SEQ ID NO: 11
20	OLIGO 12 = SEQ ID NO: 12
	OLIGO 13 = SEQ ID NO: 13
	OLIGO 14 = SEQ ID NO: 14
	OLIGO 15 = SEQ ID NO: 15
	OLIGO 16 = SEQ ID NO: 16
25	OLIGO 17 = SEQ ID NO: 17
	OLIGO 18 = SEQ ID NO: 18
	OLIGO 19 = SEQ ID NO: 19
	OLIGO 20 = SEQ ID NO: 20
	OLIGO 21 = SEQ ID NO: 21
30	OLIGO 22 = SEQ ID NO: 22
	OLIGO 23 = SEQ ID NO: 23
	OLIGO 24 = SEQ ID NO: 24
	OLIGO 25 = SEQ ID NO: 25
	OLIGO 26 = SEQ ID NO: 26

TABLE 2 Amino acid sequences of proteins, deduced from the cDNA constructs.

The full deduced sequence of the proteins of the Examples are given as follows:

- 5 MQ1->K196 of CR-1 is given in SEQ ID NO: 27
MR122->K196 of CR-1 is given in SEQ ID NO: 28
MQ1-S253 of CR-1 is given in SEQ ID NO: 29
The R235H mutant of MQ1-S253 of CR-1 is given in SEQ ID NO: 30
MR122-S253 of CR-1 is given in SEQ ID NO: 31.
- 10

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: SmithKline Beecham p.l.c. -, - -
- (ii) TITLE OF INVENTION: Novel Compounds
- (iii) NUMBER OF SEQUENCES: 31
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: SmithKline Beecham Corporate Patents
 - (B) STREET: Great Burgh, Yew Tree Bottom Road
 - (C) CITY: Epsom
 - (D) STATE: Surrey
 - (E) COUNTRY: England
 - (F) ZIP: KT18 5XQ
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Valentine, Jill B
 - (B) REGISTRATION NUMBER: G.A. 26758
 - (C) REFERENCE/DOCKET NUMBER: P30423
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 0737364158

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 87 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TATGCAGTGC AACGCTCCGG AATGGCTGCC GTTCGCGCGC CCGACCAACC TGA CTGATGA 60

ATTTGAGTTC CCGATCGGTA CCTACCT 87

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 93 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CGTAGTTCAG GTAGGTACCG ATCGGGA ACT CAAATTCATC AGTCAGGTTG GTCGGGCGCG 60

CGAACGGCAG CCATTCCGGA GCGTTGCACT GCA 93

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 101 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAACTACGAA TGCCGCCCGG GTTATAGCGG CCGCCCGTTT TCTATCATCT GCCTGAAAAA 60

CTCTGTCTGG ACTGGTGCTA AGGACCGTTG CCGACGTAAA T

101

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 101 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ACGACAAGAT TTACGTCGGC AACGGTCCTT AGCACCAGTC CAGACAGAGT TTTTCAGGCA 60

GATGATAGAA AACGGGCGGC CGCTATAACC CGGGCGGCAT T 101

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 101 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTTGTCGTAA TCCGCCAGAT CCGGTTAACG GCATGGTGCA TGTGATCAAA GGCATCCAGT 60

TCGGTTCCCA AATTAAATAT TCTTGTAATA AAGGTTACCG T 101

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 101 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCAATCAGAC GGTAACCTTT AGTACAAGAA TATTTAATTT GGGAACCGAA CTGGATGCCT 60

TTGATCACAT GCACCATGCC GTTAACCGGA TCTGGCGGAT T 101

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 94 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTGATTGGTT CCTCCAGCGC TACATGCATC ATCTCTGGTG ATACTGTCAT TTGGGATAAT 60

GAAACACCGA TTTGTGACCG AATTCAGTAA TAAA 94

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 90 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGCTTTTATT ACTGAATTCG GTCACAAATC GGTGTTTCAT TATCCCAAAT GACAGTATCA 60

CCAGAGATGA TGCATGTAGC GCTGGAGGAA 90

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 72 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TATGCGAATT CCGTGTGGTC TGCCGCCGAC CATCACCAAC GGTGATTTC TCTCTACCAA 60

TCGCGAGAAT TT 72

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 78 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CATAGTGAAG ATTCTCGCGA TTGGTAGAGA TGAAATCACC GTTGGTGATG GTCGGCGGCA 60

GACCACACGG AATTCGCA 78

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 85 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TCACTATGGT TCTGTGGTGA CCTACCGCTG CAATCCGGGT AGCGGTGGTC GTAAGGTGTT 60

TGAGCTCGTG GGTGAGCCGT CCATC 85

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 85 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTGCAGTAGA TGGACGGCTC ACCCAGGAGC TCAAACACCT TACGACCACC GCTACCCGGA 60

TTGCAGCGGT AGGTCACCAC AGAAC 85

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 79 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TACTGCACTA GTAATGACGA TCAAGTGGGC ATCTGGAGCG GCCCGGCACC GCAGTGCATC 60

ATCCCGAACA AATAATAAA 79

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 75 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AGCTTTTATT ATTTGTTTCGG GATGATGCAC TCGGGTGCCG GGCCGCTCCA GATGCCCCACT 60

TGATCGTCAT TACTA 75

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GAGACTCATA TGCAGTGCAA CGCTCCGGAA 30

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GTCAGCAAGC TTTTATTACT GAATTCGGTC 30

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ATCGTAACTA GTAACGACGA TCAAGTGGGC

30

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ATGACTAAGC TTTTATTATG AGCAGCTCGG

30

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TAACTAGTAA CGACGATCAA GTGGGCATCT GG

32

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CTAAGCTTTT ATTATGAGCA GCTCGGGAGT TCC

33

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CTAGTAACGA CGATCAAGTG GGCATCTGGA GCGGCCCGGC ACCGCAGTGC ATCATCCCCGA

60

ACAAATGCAC GCCGCCAAAT G

81

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 85 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GTTCTCCACA TTTGGCGGCG TGCATTTGTT CGGGATGATG CACTGCGGTG CCGGGCCGCT

60

CCAGATGCCC ACTTGATCGT CGTTA

85

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TGGAGAACGG TATCCTGGTA TCTGACAACC GTTCTCTGTT CTCTTTAAAC GAAGTTGTAG 60
AGTTTCGTTG TCAGCCGGGC TTTGTTATGA 90

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CGGACCTTTC ATAACAAAGC CCGGCTGACA ACGAACTCT ACAACTTCGT TTAAAGAGAA 60
CAGAGAACGG TTGTCAGATA CCAGGATACC 90

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 72 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

AAGGTCCGCG CCGTGTGAAG TGCCAGGCCT TGAACAAATG GGAGCCGGAA CTCCCGAGCT 60

GCTCATAATA AA 72

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 68 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AGCTTTTATT ATGAGCAGCT CGGGAGTTCC GGCTCCCATTT GTTCAAGGC CTGGCACTTC 60

ACACGGCG 68

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 197 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Met	Gln	Cys	Asn	Ala	Pro	Glu	Trp	Leu	Pro	Phe	Ala	Arg	Pro	Thr	Asn
1				5				10						15	

Leu	Thr	Asp	Glu	Phe	Glu	Phe	Pro	Ile	Gly	Thr	Tyr	Leu	Asn	Tyr	Glu
			20					25				30			

Cys Arg Pro Gly Tyr Ser Gly Arg Pro Phe Ser Ile Ile Cys Leu Lys
 35 40 45
 Asn Ser Val Trp Thr Gly Ala Lys Asp Arg Cys Arg Arg Lys Ser Cys
 50 55 60
 Arg Asn Pro Pro Asp Pro Val Asn Gly Met Val His Val Ile Lys Gly
 65 70 75 80
 Ile Gln Phe Gly Ser Gln Ile Lys Tyr Ser Cys Thr Lys Gly Tyr Arg
 85 90 95
 Leu Ile Gly Ser Ser Ser Ala Thr Cys Ile Ile Ser Gly Asp Thr Val
 100 105 110
 Ile Trp Asp Asn Glu Thr Pro Ile Cys Asp Arg Ile Pro Cys Gly Leu
 115 120 125
 Pro Pro Thr Ile Thr Asn Gly Asp Phe Ile Ser Thr Asn Arg Glu Asn
 130 135 140
 Phe His Tyr Gly Ser Val Val Thr Tyr Arg Cys Asn Pro Gly Ser Gly
 145 150 155 160
 Gly Arg Lys Val Phe Glu Leu Val Gly Glu Pro Ser Ile Tyr Cys Thr
 165 170 175
 Ser Asn Asp Asp Gln Val Gly Ile Trp Ser Gly Pro Ala Pro Gln Cys
 180 185 190
 Ile Ile Pro Asn Lys
 195

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 76 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met Arg Ile Pro Cys Gly Leu Pro Pro Thr Ile Thr Asn Gly Asp Phe
1 5 10 15

Ile Ser Thr Asn Arg Glu Asn Phe His Tyr Gly Ser Val Val Thr Tyr
20 25 30

Arg Cys Asn Pro Gly Ser Gly Gly Arg Lys Val Phe Glu Leu Val Gly
35 40 45

Glu Pro Ser Ile Tyr Cys Thr Ser Asn Asp Asp Gln Val Gly Ile Trp
50 55 60

Ser Gly Pro Ala Pro Gln Cys Ile Ile Pro Asn Lys
65 70 75

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 254 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Met Gln Cys Asn Ala Pro Glu Trp Leu Pro Phe Ala Arg Pro Thr Asn
1 5 10 15

Leu Thr Asp Glu Phe Glu Phe Pro Ile Gly Thr Tyr Leu Asn Tyr Glu
20 25 30

Cys Arg Pro Gly Tyr Ser Gly Arg Pro Phe Ser Ile Ile Cys Leu Lys
35 40 45

Asn Ser Val Trp Thr Gly Ala Lys Asp Arg Cys Arg Arg Lys Ser Cys

50	55	60
Arg Asn Pro Pro Asp Pro Val Asn Gly Met Val His Val Ile Lys Gly		
65	70	75 80
Ile Gln Phe Gly Ser Gln Ile Lys Tyr Ser Cys Thr Lys Gly Tyr Arg		
	85	90 95
Leu Ile Gly Ser Ser Ser Ala Thr Cys Ile Ile Ser Gly Asp Thr Val		
	100	105 110
Ile Trp Asp Asn Glu Thr Pro Ile Cys Asp Arg Ile Pro Cys Gly Leu		
	115	120 125
Pro Pro Thr Ile Thr Asn Gly Asp Phe Ile Ser Thr Asn Arg Glu Asn		
	130	135 140
Phe His Tyr Gly Ser Val Val Thr Tyr Arg Cys Asn Pro Gly Ser Gly		
145	150	155 160
Gly Arg Lys Val Phe Glu Leu Val Gly Glu Pro Ser Ile Tyr Cys Thr		
	165	170 175
Ser Asn Asp Asp Gln Val Gly Ile Trp Ser Gly Pro Ala Pro Gln Cys		
	180	185 190
Ile Ile Pro Asn Lys Cys Thr Pro Pro Asn Val Glu Asn Gly Ile Leu		
	195	200 205
Val Ser Asp Asn Arg Ser Leu Phe Ser Leu Asn Glu Val Val Glu Phe		
	210	215 220
Arg Cys Gln Pro Gly Phe Val Met Lys Gly Pro Arg Arg Val Lys Cys		
225	230	235 240
Gln Ala Leu Asn Lys Trp Glu Pro Glu Leu Pro Ser Cys Ser		
	245	250

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 254 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Met Gln Cys Asn Ala Pro Glu Trp Leu Pro Phe Ala Arg Pro Thr Asn
 1 5 10 15

Leu Thr Asp Glu Phe Glu Phe Pro Ile Gly Thr Tyr Leu Asn Tyr Glu
 20 25 30

Cys Arg Pro Gly Tyr Ser Gly Arg Pro Phe Ser Ile Ile Cys Leu Lys
 35 40 45

Asn Ser Val Trp Thr Gly Ala Lys Asp Arg Cys Arg Arg Lys Ser Cys
 50 55 60

Arg Asn Pro Pro Asp Pro Val Asn Gly Met Val His Val Ile Lys Gly
 65 70 75 80

Ile Gln Phe Gly Ser Gln Ile Lys Tyr Ser Cys Thr Lys Gly Tyr Arg
 85 90 95

Leu Ile Gly Ser Ser Ser Ala Thr Cys Ile Ile Ser Gly Asp Thr Val
 100 105 110

Ile Trp Asp Asn Glu Thr Pro Ile Cys Asp Arg Ile Pro Cys Gly Leu
 115 120 125

Pro Pro Thr Ile Thr Asn Gly Asp Phe Ile Ser Thr Asn Arg Glu Asn
 130 135 140

Phe His Tyr Gly Ser Val Val Thr Tyr Arg Cys Asn Pro Gly Ser Gly
 145 150 155 160

Gly Arg Lys Val Phe Glu Leu Val Gly Glu Pro Ser Ile Tyr Cys Thr
 165 170 175

Ser Asn Asp Asp Gln Val Gly Ile Trp Ser Gly Pro Ala Pro Gln Cys
 180 185 190

Ile Ile Pro Asn Lys Cys Thr Pro Pro Asn Val Glu Asn Gly Ile Leu
 195 200 205

Val Ser Asp Asn Arg Ser Leu Phe Ser Leu Asn Glu Val Val Glu Phe
 210 215 220

Arg Cys Gln Pro Gly Phe Val Met Lys Gly Pro His Arg Val Lys Cys
 225 230 235 240

Gln Ala Leu Asn Lys Trp Glu Pro Glu Leu Pro Ser Cys Ser
 245 250

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 133 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Met Arg Ile Pro Cys Gly Leu Pro Pro Thr Ile Thr Asn Gly Asp Phe
 1 5 10 15

Ile Ser Thr Asn Arg Glu Asn Phe His Tyr Gly Ser Val Val Thr Tyr
 20 25 30

Arg Cys Asn Pro Gly Ser Gly Gly Arg Lys Val Phe Glu Leu Val Gly
 35 40 45

Glu Pro Ser Ile Tyr Cys Thr Ser Asn Asp Asp Gln Val Gly Ile Trp
 50 55 60

Ser Gly Pro Ala Pro Gln Cys Ile Ile Pro Asn Lys Cys Thr Pro Pro
 65 70 75 80

Asn Val Glu Asn Gly Ile Leu Val Ser Asp Asn Arg Ser Leu Phe Ser
85 90 95

Leu Asn Glu Val Val Glu Phe Arg Cys Gln Pro Gly Phe Val Met Lys
100 105 110

Gly Pro Arg Arg Val Lys Cys Gln Ala Leu Asn Lys Trp Glu Pro Glu
115 120 125

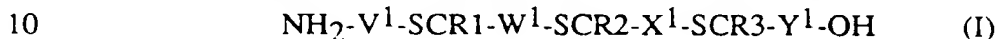
Leu Pro Ser Cys Ser
130

Claims

1. A soluble polypeptide comprising, in sequence, one to four short consensus repeats (SCR) selected from SCR 1, 2, 3 and 4 of long homologous repeat A (LHR-A) as the only structurally and functionally intact SCR domains of CR1 and including at least SCR3.

2. A polypeptide according to claim 1 which comprises, in sequence, SCR 1, 2 and 3 of LHR-A as the only structurally and functionally intact SCR domains of CR1.

3. A polypeptide according to claim 2 of formula (I):

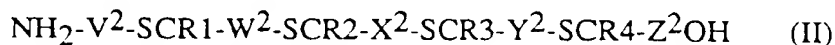


in which SCR1 represents residues 2-58 of mature CR1, SCR2 represents residues 63-120 of mature CR1, SCR3 represents residues 125-191 of mature CR1, and V^1 , W^1 , X^1 and Y^1 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

4. A polypeptide according to claim 3 in which W^1 , X^1 and Y^1 represent residues 59-62, 121-124 and 192-196, respectively, of mature CR1 and V^1 represents residue 1 of mature CR1 optionally linked via its N-terminus to methionine.

5. A polypeptide according to claim 1 which comprises, in sequence, SCR 1, 2, 3 and 4 of LHR-A as the only structurally and functionally intact SCR domains of CR1.

6. A polypeptide according to claim 5 of formula (II):



in which SCR1, SCR2 and SCR3 are as defined in claim 3, SCR4 represents residues 197-252 of mature CR1 and V^2 , W^2 , X^2 , Y^2 and Z^2 represents bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

7. A polypeptide according to claim 6 in which W^2 , X^2 , Y^2 and Z^2 represent residues 59-62, 121-124, 192-196, and residues 253 respectively, of mature CR1 and V^2 represents residue 1 of mature CR1 optionally linked via its N-terminus to methionine.

8. A polypeptide according to claim 7 in which arginine 235 is replaced by histidine.

9. A polypeptide according to claim 1 of formula (III):



in which SCR3 is as defined in claim 3 and X^3 and Y^3 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

10. A polypeptide according to claim 9 in which X^3 represents amino acids 122-124 of mature CR1 optionally linked to methionine at its N-terminus and Y^4 represents amino acids 192-196 of mature CR1.
11. A polypeptide according to claim 1 of formula (IV):
- 5
$$NH_2-X^4-SCR3-Y^4-SCR4-Z^4-OH \quad (IV)$$
- in which SCR3 and SCR4 are as defined in claim 6 and X^4 , Y^4 and Z^4 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.
12. A polypeptide according to claim 11 in which X^4 represents amino acids 122-124 of mature CR1 optionally linked to methionine at its N-terminus and Y^4 and Z^4 represent amino acids 192-196 and 253 respectively of mature CR1.
13. A polypeptide having the amino acid sequence given in SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30 or SEQ ID NO:31.
14. A DNA polymer comprising a nucleotide sequence that encodes the
- 15 polypeptide according to claim 1.
15. A replicable expression vector capable, in a host cell, of expressing the DNA polymer of claim 14.
16. A host cell transformed with the replicable expression vector of claim 15.
17. A process for preparing a CR1 polypeptide according to claim 1 which
- 20 process comprises expressing DNA encoding said polypeptide in a recombinant host cell and recovering the product.
18. A pharmaceutical composition comprising a therapeutically effective amount of a polypeptide according to claim 1 and a pharmaceutically acceptable carrier or excipient.
- 25 19. A polypeptide according to claim 1 for use as an active therapeutic substance.
20. A method of treating a disease or disorder associated with inflammation or inappropriate complement activation comprising administering to a subject in need of such treatment a therapeutically effective amount of a polypeptide according to claim 1.
- 30 21. The use of a polypeptide of claim 1 in the manufacture of a medicament for the treatment of a disease or disorder associated with inflammation or inappropriate complement activation.

Figure 1

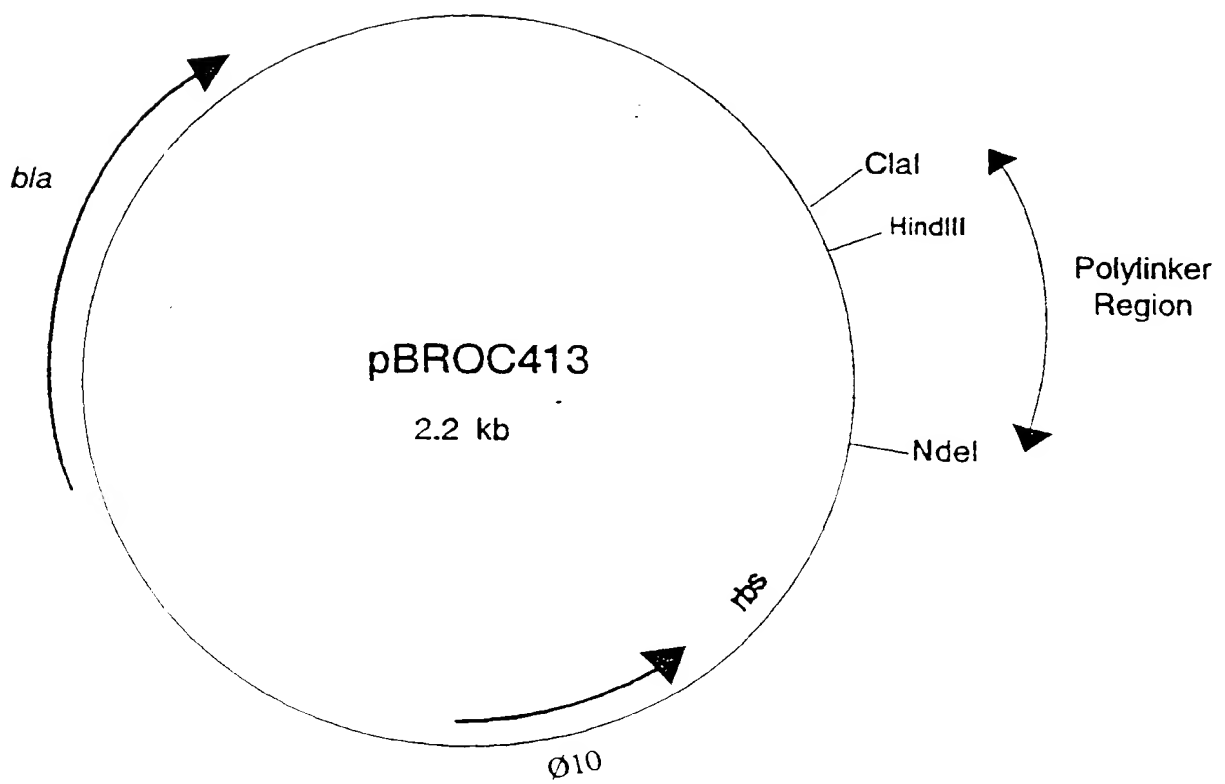


Figure 2

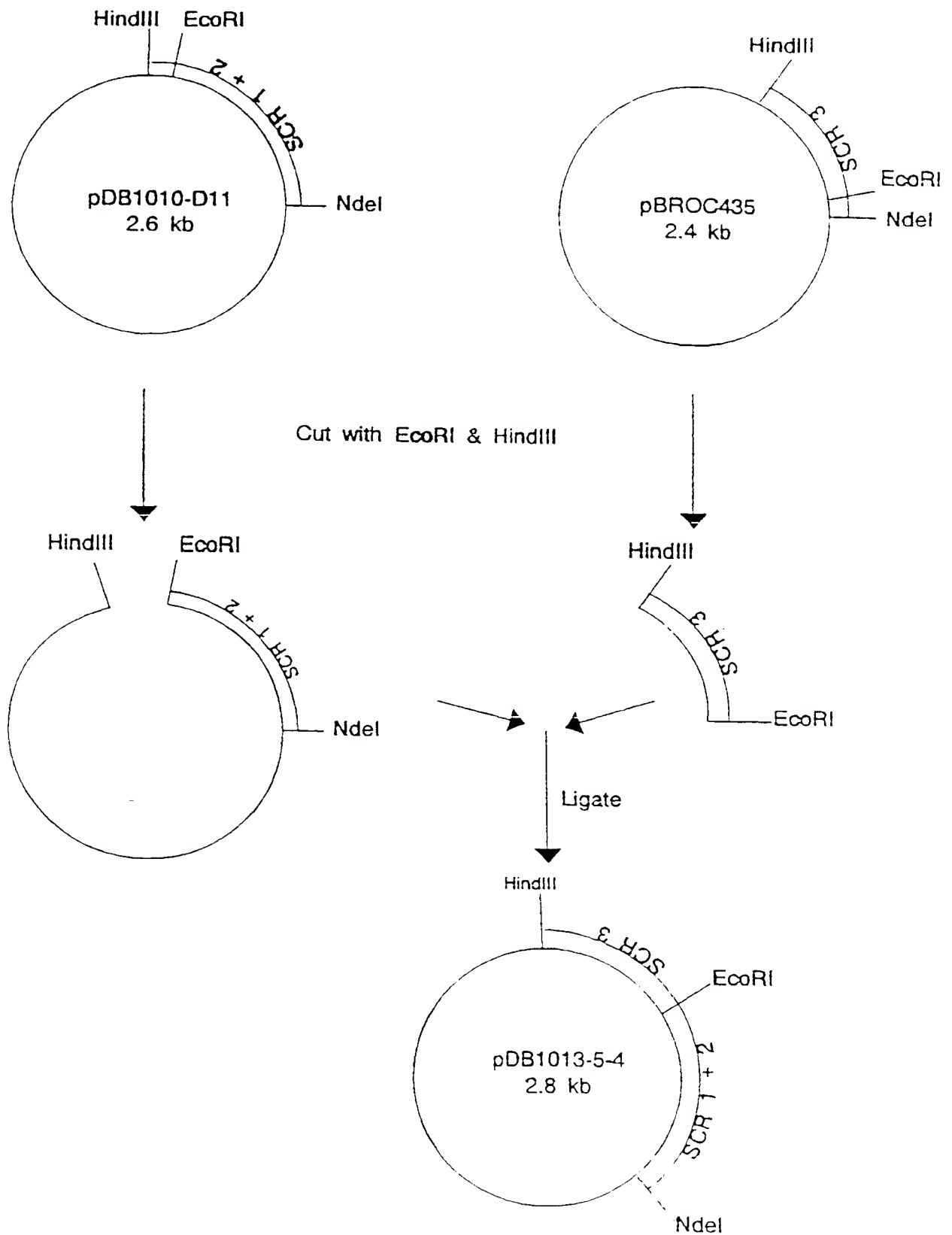
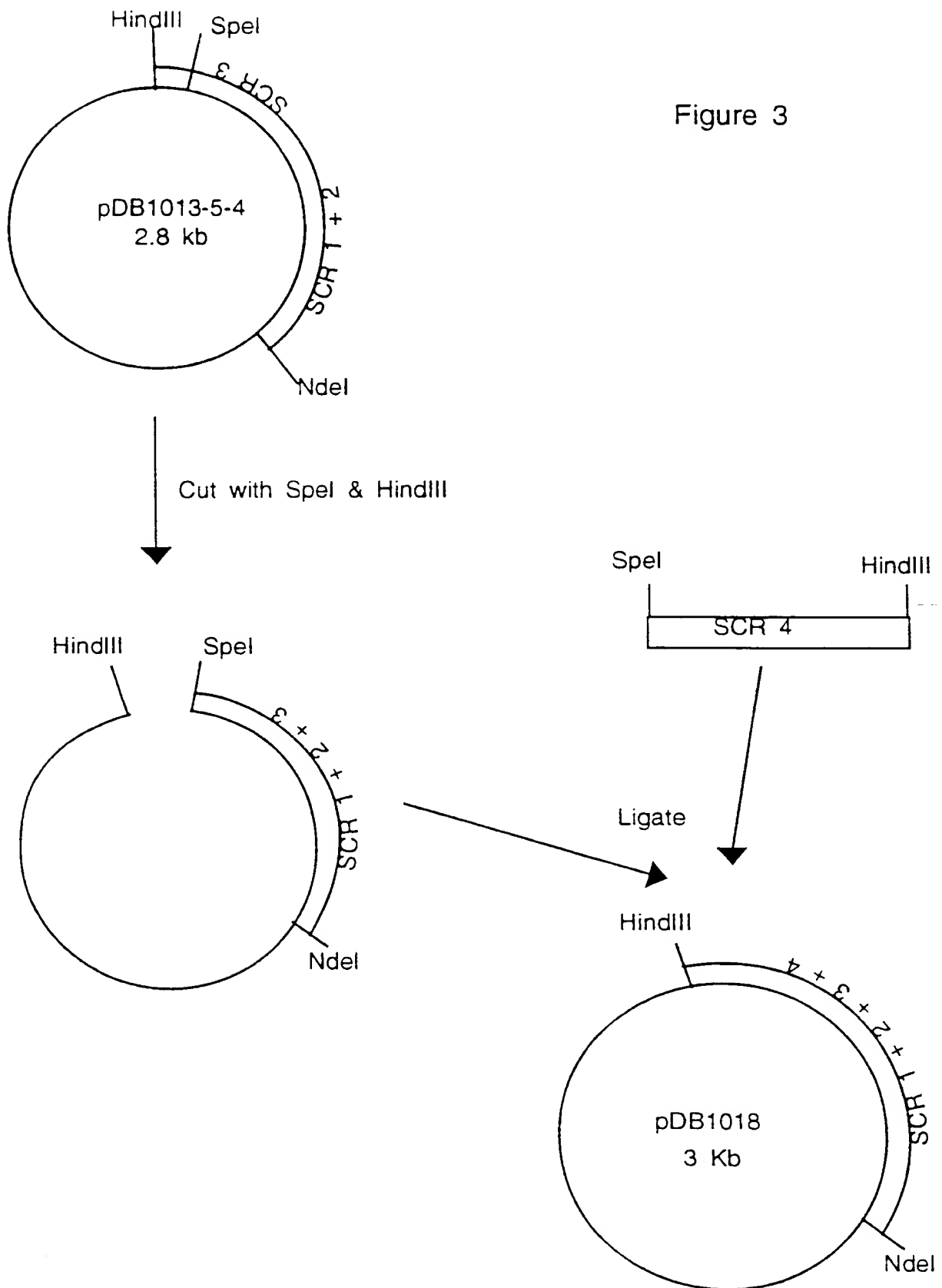


Figure 3



I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)⁶

According to International Patent Classification (IPC) or to both National Classification and IPC

Int.Cl. 5 C12N15/12; A61K37/02; C07K13/00; C12N1/21
//(C12N1/21, C12R1:19)**II. FIELDS SEARCHED**

in Europe

Minimum Documentation Searched⁷

Classification System

Classification Symbols

Int.Cl. 5

Patent doc.
not in search

C07K ;

C12N ;

A61K

Documentation Searched other than Minimum Documentation
to the Extent that such Documents are Included in the Fields Searched⁸**III. DOCUMENTS CONSIDERED TO BE RELEVANT⁹**

Category ¹⁰	Citation of Document, ¹¹ with indication, where appropriate, of the relevant passages ¹²	Relevant to Claim No. ¹³
A	WO,A,8 909 220 (THE JOHNS HOPKINS UNIVERSITY ET AL.) 5 October 1989 cited in the application see pages 32-38, chapters 5.5 and 5.6 see example 11.2.4.3 see figure 20 ---	1-21
A	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA vol. 88, no. 10, 15 May 1991, WASHINGTON US pages 4353 - 4357 KRYCH, M. ET AL. 'Sites within the complement C3b/C4b receptor important for the specificity of ligand binding' cited in the application see the whole document --- -/-	1-21

¹⁰ Special categories of cited documents : ¹⁰"A" document defining the general state of the art which is not
considered to be of particular relevance"E" earlier document but published on or after the international
filing date"L" document which may throw doubts on priority claim(s) or
which is cited to establish the publication date of another
citation or other special reason (as specified)"O" document referring to an oral disclosure, use, exhibition or
other means"P" document published prior to the international filing date but
later than the priority date claimed"T" later document published after the international filing date
or priority date and not in conflict with the application but
cited to understand the principle or theory underlying the
invention"X" document of particular relevance; the claimed invention
cannot be considered novel or cannot be considered to
involve an inventive step"Y" document of particular relevance; the claimed invention
cannot be considered to involve an inventive step when the
document is combined with one or more other such docu-
ments, such combination being obvious to a person skilled
in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search

13 OCTOBER 1993

Date of Mailing of this International Search Report

22. 10. 93

International Searching Authority

EUROPEAN PATENT OFFICE

Signature of Authorized Officer

ANDRES S.M.

III. DOCUMENTS CONSIDERED TO BE RELEVANT

(CONTINUED FROM THE SECOND SHEET)

Category °	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
A	<p>BIOLOGICAL ABSTRACTS vol. 93 , 1992, Philadelphia, PA, US; abstract no. 53233, KALLI, K. ET AL. 'Mapping of the C3b-binding site of CR1 and construction of a (CR1)2-F(ab)'12 chimeric complement inhibitor' see abstract & J.EXP.MED., 174(6), pp.1451-60; 1991 -----</p>	1-21

INTERNATIONAL SEARCH REPORT

International application No.

PCT/GB93/01282

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claim 20 is directed to a method of treatment of the human/animal body the search has been carried out and based on the alleged effects of the compound/composition

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

GB 9301282
SA 75972

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information. 13/10/93

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-8909220	05-10-89	AU-A- 3539489	16-10-89
		EP-A- 0411031	06-02-91
		JP-T- 4501502	19-03-92
		US-A- 5212071	18-05-93
